

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 24, 2004, 07:39:56 ; Search time 4765.54 Seconds
(without alignments)
8565.996 Million cell updates/sec

Title: US-09-743-823-21

Perfect score: 1367

Sequence: 1 taaataatattacattaa.....ctacacacactagaagaata 1367

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 segs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_estc:*
9: gb_estl:*
10: gb_est2:*
11: gb_est3:*
12: gb_est4:*
13: gb_est5:*
14: gb_estfun:*
15: em_estom:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gssl:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------|
| 1 | 83.8 | 6.1 | 1200 | 13 | BX437758 |
| 2 | 80 | 5.9 | 1200 | 13 | BX415878 |
| 3 | 79.8 | 5.8 | 887 | 13 | BX441520 |
| 4 | 79.4 | 5.8 | 1201 | 13 | BX446296 |

| | | | | | |
|----|------|-----|------|----|----------|
| 5 | 79 | 5.8 | 1165 | 13 | BX338369 |
| 6 | 77.6 | 5.7 | 987 | 29 | CNS014PQ |
| 7 | 76.6 | 5.6 | 1201 | 13 | BX334085 |
| 8 | 76.4 | 5.6 | 887 | 13 | BX441520 |
| 9 | 75.4 | 5.5 | 1101 | 29 | CNS00K35 |
| 10 | 75.2 | 5.5 | 1101 | 29 | CNS0039G |
| 11 | 75.2 | 5.5 | 1201 | 13 | BX446296 |
| 12 | 75.2 | 5.5 | 1391 | 29 | CG754863 |
| 13 | 75 | 5.5 | 905 | 29 | CNS00KHX |
| 14 | 74.4 | 5.4 | 1101 | 29 | CNS00EVL |
| 15 | 74.4 | 5.4 | 1200 | 13 | BX415878 |
| 16 | 74.2 | 5.4 | 1201 | 9 | AL536104 |
| 17 | 73.8 | 5.4 | 1201 | 9 | AL536104 |
| 18 | 73.6 | 5.4 | 1201 | 13 | BX437779 |
| 19 | 73.4 | 5.4 | 888 | 28 | AZ549422 |
| 20 | 73.4 | 5.4 | 928 | 29 | CNS00DKY |
| 21 | 73.4 | 5.4 | 1201 | 13 | BX421282 |
| 22 | 73.2 | 5.4 | 681 | 13 | BX378802 |
| 23 | 73 | 5.3 | 928 | 29 | CNS00DKY |
| 24 | 73 | 5.3 | 1200 | 13 | BX437758 |
| 25 | 72.8 | 5.3 | 1098 | 13 | BX377526 |
| 26 | 72.6 | 5.3 | 911 | 13 | BX446782 |
| 27 | 72.6 | 5.3 | 1162 | 9 | AL513719 |
| 28 | 72.6 | 5.3 | 1201 | 13 | BX446356 |
| 29 | 72.4 | 5.3 | 1092 | 29 | CNS020K7 |
| 30 | 72.4 | 5.3 | 1201 | 9 | AL561317 |
| 31 | 72.4 | 5.3 | 1201 | 13 | BX440226 |
| 32 | 72.2 | 5.3 | 1201 | 13 | BX443774 |
| 33 | 72 | 5.3 | 986 | 9 | AL514691 |
| 34 | 72 | 5.3 | 1201 | 9 | AL532464 |
| 35 | 71.8 | 5.3 | 983 | 13 | BX456702 |
| 36 | 71.8 | 5.3 | 1201 | 13 | BX406178 |
| 37 | 71.4 | 5.2 | 633 | 13 | BX403519 |
| 38 | 71.4 | 5.2 | 1101 | 29 | CNS0022U |
| 39 | 71.4 | 5.2 | 1201 | 13 | BX360615 |
| 40 | 71.2 | 5.2 | 1101 | 29 | CNS00EQL |
| 41 | 71.2 | 5.2 | 1101 | 29 | CNS00EVL |
| 42 | 70.8 | 5.2 | 1098 | 13 | BX377526 |
| 43 | 70.8 | 5.2 | 1104 | 9 | AL513809 |
| 44 | 70.6 | 5.2 | 1056 | 13 | BX415058 |
| 45 | 70.6 | 5.2 | 1201 | 9 | AL532464 |

ALIGNMENTS

RESULT 1
BX437758
LOCUS
DEFINITION
BX437758 Homo sapiens THYMUS Homo sapiens cDNA clone CS0CAP008YB01
5-PRIME, mRNA sequence.
ACCESSION
BX437758
VERSION
BX437758.1
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1200)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
COMMENT
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0CAP008CA01QPL.
Location/Qualifiers
1. .1200
/organism="Homo sapiens"

EST 15-MAY-2003
linear
mRNA
1200 bp
Homo sapiens THYMUS Homo sapiens cDNA clone CS0CAP008YB01


```

Db      1154 WATTAATAATWATTTTWTAMMTWAAAAAAT 1184

RESULT 3
BX441520/c
LOCUS      BX441520
DEFINITION BX441520 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
            CS0DF018YJ11 3-PRIME, mRNA sequence.
ACCESSION  BX441520
VERSION     BX441520
KEYWORDS    EST.
SOURCE      BX441520.1 GI:30787976
            Homo sapiens (human)
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 887)
AUTHORS     Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE       Full-length cDNA libraries and normalization
JOURNAL     Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. This sequence belongs to sequence cluster 7316.r
            Contact: Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CS0DF018CE06NP1.
FEATURES    Location/Qualifiers
             1..887
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="CS0DF018YJ11"
                /tissue types="FETAL BRAIN"
                /dev stage="fetal"
                /clone lib="Homo sapiens FETAL BRAIN"
                /notes="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
                    was primed with a NotI-oligo(dT) primer. Five prime end
                    enriched, double-strand cDNA was digested with Not I and
                    cloned into the Not I and EcoRV sites of the pCMVSPORT 6
                    vector. Library was not normalized."
ORIGIN
Query Match      5.8%; Score 79.8; DB 13; Length 887;
Best Local Similarity 42.6%; Pred. No. 0.0019;
Matches 141; Conservative 44; Mismatches 146; Indels 0; Gaps 0;

Qy      1 TAAATAATCTATACATTAAAAAATTTGATTTAAAAATTTTACAAATTCATGATTTATT 60
Db      1 TAAATAATCTATACATTAAAAAATTTGATTTAAAAATTTTACAAATTCATGATTTATT 60

Qy      857 TAAATATTTTATTTTATATATATATATATATATATATATATATATATATATATAT 798
Db      857 TAAATATTTTATTTTATATATATATATATATATATATATATATATATATATATAT 798

Qy      61 TTTTTCACGAAATCCGTTAATTTGTTAAATATACCACTAATTTTAAATTTTA 120
Db      61 TTTTTCACGAAATCCGTTAATTTGTTAAATATACCACTAATTTTAAATTTTA 120

Qy      797 AATTTTWTWTAATAATTTWTWAAATTTWTWAAATATATATATTTTAAATTTTAA 738
Db      797 AATTTTWTWTAATAATTTWTWAAATTTWTWAAATATATATATTTTAAATTTTAA 738

Qy      121 TTTTAAAGCAATTAAGCATGTTTGATATAATATATATATATATATATATATATATAT 180
Db      121 TTTTAAAGCAATTAAGCATGTTTGATATAATATATATATATATATATATATATATAT 180

Qy      737 WTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 678
Db      737 WTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 678

Qy      181 GTATAAAGTTGATGATCGCGTGGTGGTAGATATTTTAGTTCTAGGTCGAATTCGAAGTT 240
Db      181 GTATAAAGTTGATGATCGCGTGGTGGTAGATATTTTAGTTCTAGGTCGAATTCGAAGTT 240

Qy      677 WTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT 618
Db      677 WTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT 618

Qy      241 GGTTTAGACATTTAGCCTTATCTTTTCTTAAACCAAAATAAATGTAATGGAACCTT 300
Db      241 GGTTTAGACATTTAGCCTTATCTTTTCTTAAACCAAAATAAATGTAATGGAACCTT 300

Qy      617 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 558
Db      617 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 558

Qy      301 TAGGAAAAAAGAAATCAAAATTCGAAACA 331
Db      301 TAGGAAAAAAGAAATCAAAATTCGAAACA 331

Qy      557 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 527
Db      557 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 527

RESULT 4
BX446296
LOCUS      BX446296
DEFINITION BX446296 Homo sapiens NEUROBLASTOMA Homo sapiens cDNA clone
            XCL0BB001ZA03 3-PRIME, mRNA sequence.
ACCESSION  BX446296
VERSION     BX446296
KEYWORDS    EST.
SOURCE      BX446296.1 GI:30784407
            Homo sapiens (human)
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 1201)
AUTHORS     Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE       Full-length cDNA libraries and normalization
JOURNAL     Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. Contact: Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
            Faraday Avenue Genoscope sequence ID : XCL0BB001ZA03FP1.
FEATURES    Location/Qualifiers
             1..1201
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="XCL0BB001ZA03"
                /tissue type="NEUROBLASTOMA"
                /clone lib="Homo sapiens NEUROBLASTOMA"
                /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
                    with a NotI-oligo(dT) primer. Five prime end enriched,
                    double-strand cDNA was digested with Not I and cloned into
                    the Not I and EcoRV sites of the pCMVSPORT 6 vector.
                    Library was not normalized."
ORIGIN
Query Match      5.8%; Score 79.4; DB 13; Length 1201;
Best Local Similarity 38.8%; Pred. No. 0.0018;
Matches 143; Conservative 60; Mismatches 166; Indels 0; Gaps 0;

Qy      603 TTGGTCTTTGCTTTTGGCTTTTGTTCATTTTTCACGTGATTTTGTGGTGAACACGCG 662
Db      127 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 186

Qy      663 TACTTGAACTTATTAATAATTAATAATTAATAATTAATAATTAATAATTAATAATTAATA 722
Db      187 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 246

Qy      723 TATAATATATAGGTTTGAATGCCAATTTTAAATAAAGATAATAATAATAATAATAATAAT 782
Db      247 WTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 306

Qy      783 AAAAAATCGAAAAATGAAATGTAAAAAATTTGAGGGGGCAATAATAAATATGAAATCTAT 842
Db      307 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 366

Qy      843 TATTTAAATTTCCATAGAAATCTATTTCTCTAGTTAATATGAGTACGAGTTGGGA 902
Db      367 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 426

Qy      903 GATACACGAAAAATGTCATGAAACAGTTGCATCTAGGAAATTAATCTAGTAGGGGATAG 962
Db      427 RAARARAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 486

Qy      963 CAAGCAAA 971
Db      487 AAAAAAAA 495

```

```

BX338369
LOCUS
DEFINITION
BX338369 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
Clone CS0D1058YF14 5-PRIME, mRNA sequence.
ACCESSION
VERSION
BX338369
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
REFERENCE
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1165)
AUTHORS
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE
Full-length cDNA libraries and normalization
JOURNAL
Unpublished (2001)
COMMENT
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/InvitroGen Corporation 1600
Paradise Avenue Genoscope sequence ID : CS0D1058DC07QPI.
FEATURES
Location/Qualifiers
source
1..1165
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1058YF14"
/issue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN
Query Match 5.8%; Score 79; DB 13; Length 1165;
Best Local Similarity 39.3%; Pred. No. 0.0022;
Matches 130; Conservative 58; Mismatches 143; Indels 0; Gaps 0;
QY 1 TAAATAAATCTATACATTAATAAATTTGATTTTAAATTTTAAATTTTGAATTCATGATTTTAT 60
|||||
DB 813 TAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 872
QY 61 TTTTITACAGAAATCCGTTAATATGTGTAATAATATACCAATTAATTAATAAATTTA 120
|||||
DB 873 TAAATAATATTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 932
QY 121 TTTTACAGCAATTAAGCATTTGTATATAATATATATATATATATATATATATATATATAT 180
|||||
DB 933 TTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 992
QY 181 GTATAAGTTGATGCGGTGCTAGATTAATTTTGTAGTTCTAGTTTCAAGTCAAGTT 240
|||||
DB 993 TATWATTTTWTWATTTTWTWATTTTWTWATTTTWTWATTTTWTWATTTTWTWATTTTWTW 1052
QY 241 GGTITAGACATTTAGCTTATCTTTTCTTACCAATAATAATTAATTAATTAATTAATTAAT 300
DB 1053 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1112
QY 301 TAGGAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 331
DB 1113 WTTTATWTTTWTWATTTTWTWATTTTWTWATTTTWTWATTTTWTWATTTTWTWATTTTWT 1143
RESULT 6
CNS014PQ
LOCUS
DEFINITION
CNS014PQ Drosophila melanogaster genome survey sequence SP6 end of BAC
BACN12P22 of DrosBAC library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION
AL104456

```

```

AL104456.1 GI:5616067
GSS.
Drosophila melanogaster (fruit fly)
ORGANISM
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Insecta; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 987)
REFERENCE
Genoscope.
Direct Submission
Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
- Web : www.genoscope.cns.fr
COMMENT
Determination of this BAC-end sequence was carried out as part of a
collaboration with the European Drosophila Genome Project (EDGP) -
http://www.edgp.ebi.ac.uk/. This Drosophila melanogaster BAC
library (Dros BAC) was made by Alain Billaud at CEPH (Centre
d'Etude du Polymorphisme Humain) with funding provided by a MRC
project grant. The DNA was prepared from embryos by Alain Bucheton
and Genevieve Payan. It has been constructed in the vector
pBelOBAC11.
FEATURES
Location/Qualifiers
source
1..987
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACN12P22"
/clone_lib="DrosBAC"
/plasmid="pBelOBAC11"
/note="end : SP6"
ORIGIN
Query Match 5.7%; Score 77.6; DB 29; Length 987;
Best Local Similarity 38.3%; Pred. No. 0.004;
Matches 79; Conservative 62; Mismatches 65; Indels 0; Gaps 0;
QY 564 ACTTGAACCTATTAATAATACATATTTTATATAGTTTCACCTCTTATATATATCTCAT 723
DB 768 AATTAAATTTTWTWTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 827
QY 724 ATAATATATAGGTTAGATGCAATTTTAAATAAATAAATAAATAAATAAATAAATAAATAA 783
DB 828 ANAKWTTTWTTTTWTWTTTATTAATTTTAAATAATTTTAAATAATTTTAAATAATTTTAA 887
QY 784 AATTCGAAAAATCAATGTAAAAATTTGAGGGGACAAATAAATAATTAATTAATTAATTAAT 843
DB 888 ATWAAAAAATWAAAAAATWAAAAAATWAAAAAATWAAAAAATWAAAAAATWAAAAAATW 947
QY 844 ATTTAAATTTTCCATTAGATTTCTAT 869
DB 948 WWWWWWTTTTTTTTTTTAKWTKNK 973
RESULT 7
BX334085
LOCUS
DEFINITION
BX334085 Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED Homo sapiens
cDNA clone CS0DD06YL06 5-PRIME, mRNA sequence.
ACCESSION
VERSION
BX334085
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
REFERENCE
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1201)
AUTHORS
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE
Full-length cDNA libraries and normalization
JOURNAL
Unpublished (2001)
COMMENT
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

```

Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 6562.r For more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DD06DF03QP&cluster=6562.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Paradise Avenue Genoscope sequence ID : CS0DD06DF03QP1.

FEATURES source
1. .1201
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DD06DF03QP1"
/tissue_type="NEUROBLASTOMA COT 50-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and EcoRV sites of the pcMVSPORT 6 vector. Library was normalized."

ORIGIN
Query Match 5.6%; Score 76.6; DB 13; Length 1201;
Best Local Similarity 45.5%; Pred. No. 0.0051;
Matches 145; Conservative 29; Mismatches 145; Indels 0; Gaps 0;

QY 2 AAAATAATCTATACATTAATAAAATTCATTTTAAATTTAGAAATTCATGATTTATTT 61
Db AAAAAAAAAAAAAAAAAAAATTTATATAAATGCGSAAATTAATAAATAAAAWW 931
QY 62 TTTTACCAGAAATCGTTAATATGTTAAATATTACCACTAATTAATTAATTTAT 121
Db WAAAAAAAAAAAAAAAAAAATTTTAAAAAAAAAAAAAAAAAAATTTTAAAAATTTT 991
QY 122 TTTAAGGCAATTAAGCATGTTTGTATAATATATATATATATATATATATATATAT 181
Db TTTTAAAAAAAAATTTTATTTTAAAAAAAAAAATTTTAAAAAAAAAAAAAAAA 1051
QY 182 TATAAGTGAATGAGCGGCGGTAGATATTTTATGTTCTAGTTCGATGCAAGTTG 241
Db 1052 WAAAAATTTWWTT 1111
QY 242 GTTTAGCATTTAGCCTTATCTTTTCTAACCAATTAATGTAATGGAACCTTT 301
Db 1112 NTTT 1171
QY 302 AGGAAAAAAGAAATCAA 320
Db 1172 AAAAAAAAAAAATWTWAMAA 1190

RESULT 8
BX441520 887 bp mRNA linear EST 15-MAY-2003
LOCUS BX441520 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
DEFINITION CS0DF018YU11 3-PRIME, mRNA sequence.
ACCESSION BX441520
VERSION BX441520.1 GI:30787976
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 887)
Li.W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seque@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 7316.r

Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Paradise Avenue Genoscope sequence ID : CS0DF018CE06NPI.

FEATURES source
1. .887
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DF018YU11"
/tissue_type="FETAL BRAIN"
/dev_stage="fetal"
/clone_lib="Homo sapiens FETAL BRAIN"
/note="Organ: Brain; Vector: pcMVSPORT_6; 1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and EcoRV sites of the pcMVSPORT 6 vector. Library was not normalized."

ORIGIN
Query Match 5.6%; Score 76.4; DB 13; Length 887;
Best Local Similarity 42.5%; Pred. No. 0.0066;
Matches 181; Conservative 44; Mismatches 200; Indels 1; Gaps 1;

QY 550 TTATGTTAGGTAGTGTGTTGCAACAGTTGAAGCGCATCTCAGGAGCTACTTGGTCT 609
Db TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 485
QY 610 TCGTTTTCGCTCTTTGTTCAATTTTTCACGATTTTGTGTTGCAACACCGTACTTGA 669
Db 486 KXTTT 545
QY 670 AACTTATTAATAATACATAATTTTATAAGTTTCTTCTTATATAATCTCATATAA 729
Db 546 TTT 605
QY 730 TATAGGTTTGAATGCCAATTTTAAAAAAGAAATAAAAAATATAAGATAAAATCG 789
Db 606 TAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 665
QY 790 AAAAAATGAATGTAAAAAATTTGAGGGGCAAAATAAAATATGAAGTCTATTATTA 849
Db 666 TAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 725
QY 850 ATTTCCTATTGAATCTTATTTTCTTCTAGTTA-ATATGAGTTCAGCAGTTGGAGATACA 908
Db 726 TWTWTTTWTAAATTTTWTATTTTWTATTTTWTATTTTWTATTTTWTATTTTWT 785
QY 909 CGAAAAATGTCATGAACACAGTTCATGTCAGGAAATTAATGATGAGGATAGCAAGAC 968
Db 786 TAWAWAAAAATTTATAAATAATTTTWTATTTTWTATTTTWTATTTTWTATTTTWT 845
QY 969 AAAAAAT 974
Db 846 AAAAAAT 851

RESULT 9
CNS00KBS 1101 bp DNA linear GSS 03-JUN-1999
LOCUS CNS00KBS
DEFINITION Drosophila melanogaster genome survey sequence T7 end of BAC:
BACR39P05 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION AL077453.1 GI:4956930
VERSION AL077453
KEYWORDS GSS.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 1101)
Genoscope.
Direct Submission

JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
 Web : www.genoscope.cns.fr

COMMENT Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammeter in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the BDGP's pl and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES Location/Qualifiers
 source 1..1101
 /organism="Drosophila melanogaster"
 /mol_type="genomic DNA"
 /db_xref="taxon:7227"
 /clone="BACR39P05"
 /clone_lib="RPCI-98"
 /note="end : T7"

ORIGIN
 Query Match 5.5%; Score 75.4; DB 29; Length 1101;
 Best Local Similarity 48.2%; Pred. No. 0.0084;
 Matches 149; Conservative 10; Mismatches 149; Indels 0; Gaps 0;

QY 666 TTGAACATTATTAATAATACATAAATTTATTAAGTTTCATCTTCTTATATAATACATCAT 725
 Db 183 TTTAAAAATWATAAAAAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 242
 QY 726 AATATATAGGTTTACGATCCCAATTTTAAAGAAAGATTAATAATAATAATTAATAA 785
 Db 243 AAAAAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAA 302
 QY 786 ATCGAAAAATGAATGTAAAAATTTGAGGGGACAAATAATAATAATAATAATAATAATA 845
 Db 303 ATATAAAAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 362
 QY 846 TTAATTTTCCATAGAAATCTATTTTCTTAGTTAATATAGCTAGCAGTTGGAGAT 905
 Db 363 TAATATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 422
 QY 906 ACACGAAATGTCATGAAACAGTTGATGAGGAAATTAATGTTAGTAGGGATAGCAA 965
 Db 423 AAAAAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAA 482
 QY 966 GACAAAA 972
 Db 483 AAANAAA 489

RESULT 10
 LOCUS CNS0039G 1101 bp DNA linear GSS 03-JUN-1999

DEFINITION Drosophila melanogaster genome survey sequence TET3 end of BAC # BACR08K10 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.

ACCESSION AL063921
 VERSION AL063921.1 GI:4941778

KEYWORDS GSS.

SOURCE Drosophila melanogaster (fruit fly)

ORGANISM Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 Ephydroidea; Drosophilidae; Drosophila.
 1 (bases 1 to 1101)
 REFERENCE 1
 AUTHORS Genoscope.

TITLE Direct Submission
 JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
 Web : www.genoscope.cns.fr

COMMENT Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammeter in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the BDGP's pl and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES Location/Qualifiers
 source 1..1101
 /organism="Drosophila melanogaster"
 /mol_type="genomic DNA"
 /db_xref="taxon:7227"
 /clone="BACR08K10"
 /clone_lib="RPCI-98"
 /note="end : TET3"

ORIGIN
 Query Match 5.5%; Score 75.2; DB 29; Length 1101;
 Best Local Similarity 18.5%; Pred. No. 0.009;
 Matches 130; Conservative 288; Mismatches 278; Indels 8; Gaps 1;

QY 14 ACATAAAAAATTTGATTTTAAAAATTTTGAATTTTATGATTTTATTTTATTTTATACAGA 73
 Db 396 ATATAAAMWMTTITTTTTTAAATAAATAATTTAAAWAAAAAATTTAAAWAAAA 455
 QY 74 AATCCGTTATATTTGTTAAAAATATTACCAACTAATTTATTAATTTATTTTATAGGCAAT 133
 Db 456 AWAATWATTTATTAATAAAAAAATAATTTTATTTTATTTTATTTTATTTTATTTTAA 515
 QY 134 AAGCATCTTTGATAAAAATATATATTTTATATAAATCTTTTCAAAAGTATAAGTTGAT 193
 Db 516 AAAAAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAA 575
 QY 194 GATGGCGTGGTGGTAGAATTTATTTAGTTCTGATTCGAATGCAAGTTGGTTTGAACATTT 253
 Db 576 TTTTTHYTYTYTMTTTHYTHYTHYTHYTHYTHYTHYTHYTHYTHYTHYTHYTHYTHYTHY 635
 QY 254 AGCTTATCTTTTCTTACCAATAATAATGTAATGGAACCTTTAGGAAAAAAG 313
 Db 636 TAAATTTTCTTTTTHYTHYTHYTHYTHYTHYTHYTHYTHYTHYTHYTHYTHYTHYTHYTHY 695
 QY 314 A-----AATCAAAATTTGAAAAACATCATCCGGTGGAGTCGAGAGCCACACCCACGT 365
 Db 696 HTCHHCYVYVHYTHYTHYTHYTHYTHYTHYTHYTHYTHYTHYTHYTHYTHYTHYTHYTHY 755
 QY 366 GACCAACAATATAATAATGAGAGTTTGTCTCTACAGTAATGCGATACCTTTTATATCAA 425
 Db 756 TWYAWAHAMWMTTTHYTHYTHYTHYTHYTHYTHYTHYTHYTHYTHYTHYTHYTHYTHYTHY 815
 QY 426 TACTTTTCCATCTTAAATCTTGGAGATTTTGCACCGTTTAACTAATTAAGTTTATATC 485
 Db 816 WHYHTATCTMTTTHYTHYTHYTHYTHYTHYTHYTHYTHYTHYTHYTHYTHYTHYTHYTHY 875
 QY 486 CAACGGTCTAAAAAATCTTGTCTACCGTCCCTCACTTCACTTTGGCGACCCCTGAAA 545
 Db 876 CWACMTMHHWMTTTHYTHYTHYTHYTHYTHYTHYTHYTHYTHYTHYTHYTHYTHYTHYTHY 935
 QY 546 GCGGTATGTTAGTTAGTTAGTTTTCACAGAGTGAAGCGCATCATCTCAGAGGCTACTTG 605
 Db 936 HHHHWWATWMTTMTTMMCM 995

```

QY 606 GTCTGCTTTGGTCTTTGTTCAATTTTTCAGTGATTTGTTGGTGAACACGGGTAC 665
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
996 CYHYCTWHTYATWAWTAHATTAATTTATWWWHEHWAHATWWWWWATAWACTCHEW 1055
QY 666 TTGAACCTTATTAATAATACAAATTTATTAAGTTCACTTCT 709
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1056 TWYHTCTWYHYHTYHWWAWWAWHWHHMYAHYHWAHHCWYTT 1099

RESULT 11
EX446296/c 1201 bp mRNA linear EST 15-MAY-2003
LOCUS EX446296 Homo sapiens NEUROBLASTOMA Homo sapiens cDNA clone
DEFINITION XLOCBB001ZA03 3-PRIME, mRNA sequence.
ACCESSION EX446296
VERSION EX446296.1 GI:30784407
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1201)
AUTHORS Li W.B., Gruber C., Jesse J. and Polayes D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : XLOCBB001ZA03FFP1.
FEATURES
Location/Qualifiers
source 1..1201
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="XLOCBB001ZA03"
/tissue_type="NEUROBLASTOMA"
/clone_lib="Homo sapiens NEUROBLASTOMA"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NorI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."
ORIGIN
Query Match 5.5%; Score 75.2; DB 13; Length 1201;
Best Local Similarity 40.4%; Pred. No. 0.0085;
Matches 101; Conservative 53; Mismatches 96; Indels 0; Gaps 0;
QY 603 TTGCTCTGCTTTGGCTTTTCTTCTCAATTTTTCAGTGATTTTGGTGAACACGG 662
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
332 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 273
QY 563 TACTTGAACCTATTATAAATACATAATTTTATAGTTTCACCTCTTATATACTCA 722
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
272 TATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 213
QY 723 TATAATATATAGGTTTAGATGCAATTTTAAAAAAGAAATAAAAAATAAATAGAA 762
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
212 WWWWWWTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 153
QY 783 AAAATCGAAAAATGAAATGTAAAAATTTGAGGGGACAAATAAATATGAAAGTCTAT 842
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
152 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 93
QY 843 TATTTAAATTT 852
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
92 AAAAAAATTT 83

```

```

RESULT 12
CG754863 1391 bp DNA linear GSS 24-OCT-2003
LOCUS P050-2-G05.za Ppa EcoRI BAC Library Pristionchus pacificus genomic,
DEFINITION Genomic survey sequence.
ACCESSION CG754863
VERSION CG754863.1 GI:37980782
KEYWORDS GSS.
SOURCE Pristionchus pacificus
ORGANISM Pristionchus pacificus
Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
Neodiplogasteridae; Pristionchus.
REFERENCE 1 (bases 1 to 1391)
AUTHORS Srinivasan, J., Sinz, W., Jesse, T., Wiggers-Perebolte, L., Jansen, K.,
Buntjer, J., van der Meulen, M. and Sommer, R.J.
TITLE An integrated physical and genetic map of the nematode Pristionchus
pacificus
JOURNAL Mol. Genet. Genomics 269 (5), 715-722 (2003)
MEDLINE 22835951
PUBMED 12884007
COMMENT Contact: Sommer RJ
Evolutionary Biology
Max-Planck-Institute for Developmental Biology
Spemannstr. 37-39, Tuebingen D-72076, Germany
Tel: 00497071601371
Fax: 00497071601498
Email: raif.sommer@tuebingen.mpg.de
Class: BAC ends.
FEATURES
Location/Qualifiers
source 1..1391
/organism="Pristionchus pacificus"
/mol_type="genomic DNA"
/strains="California"
/db_xref="taxon:54126"
/clone_lib="Ppa EcoRI BAC library"
/note="The library was generated by a partial digest of
the genomic DNA with EcoRI and cloning into the BAC
vector."
ORIGIN
Query Match 5.5%; Score 75.2; DB 29; Length 1391;
Best Local Similarity 43.6%; Pred. No. 0.0078;
Matches 377; Conservative 0; Mismatches 476; Indels 10; Gaps 2;
QY 1 TAAATATCTATACATTAAAAAAATTTGATTTTAAAAATTTTAGAAATTCATGATTTTAT 60
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1198 TATATTAAAAATTTAAATATATATTTTAAATATATTTTAAATATATTTTAAATATTTAT 1129
QY 61 TTTTITTTACGAAATCCGTTAATATGTTAAAAATTTTACCACATAATTTATAAATTTTA 120
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1128 ATATTTTATANATAATTTATATAAATAATAAATTTAT-----ATAATTAATAATTTTA 1075
QY 121 TTTTAAGGCAATAGCAGTTTGATAAATAATATATATTTGTTATAAATACTTTTCAAAA 180
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1074 ATATATTAATAAAAAATTAATTTTAAATNTATATAATTTTAAATATATTTATATTANAA 1015
QY 181 GTATAAAGTTGATGATGGCGTGTGATATTTTGTAGTTCTAGTTTCAAGTCAAGTT 240
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1014 ATTANTATATATTTTTPAATAAATAAATAAATTTATATAAATTTAAAAATTTAAAT 955
QY 241 GGTTTAGCAATTTAGCTTTATTTCTTTTCTTAACCAATAAATGTAATGAAACCTT 300
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
954 AAAATTTATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 895
QY 301 TAGGAAAAAAGAAATCAAAATTCGAAATCATCCGTTGGAGTCGAGAGCCACACC 360
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
894 ATTATAAAAATTAATATATTTATAAAAATAAATAAATAAATAAATAAATAAATAA 835
QY 361 CACGTGACCCCAACAATTTAAAAATAAGAGTTTGCTCTACAGTAATGCGATACCTTTT 420
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
834 ATATAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 775
QY 421 TTCAATACCTTTTCCACTTCTAAAAATCTTGGAGATTTGCACCGTTAACTAATTAAGTGT 480

```


FEATURES
 source
 Faraday Avenue Genoscope sequence ID : CS0CAP008BE02Qb1.
 Location/Qualifiers
 1..1200
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0CAP008Y104"
 /tissue_type="THYMUS"
 /clone_lib="Homo sapiens THYMUS"
 /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector.
 Library was not normalized."

ORIGIN
 Query Match 5.4%; Score 74.4; DB 13; Length 1200;
 Best Local Similarity 35.3%; Pred. No. 0.011;
 Matches 308; Conservative 117; Mismatches 434; Indels 13; Gaps 3;

| | | | |
|----|------|---|------|
| QY | 23 | AATTTGATTTTAAATTTTAGAAATTCATGATTTATTTTTTTTACCAGAAATCGGTTA | 82 |
| DB | 1198 | AAAWNAWAAAAAAWATTTTWWKNTAWAAATWATTTATWATKWWTAATWTAATTT | 1139 |
| QY | 83 | ATATGTTTAAATATATPACCACTAATTTATAAATTTTATTTTAAGCAATTTAAGCATGTT | 142 |
| DB | 1138 | TTATTTTAAATAWAAWAAWAAWTTWATATTTATTTWATAAAAAATTTATWGAATATWTTT | 1079 |
| QY | 143 | TGATAAAATATATATATCTGTATAAATACCTTTTCAAAAGTATAAAGTTGATGATCGCGT | 202 |
| DB | 1078 | WAAAWATTTTAWTAAATATATATAWAAAAWTTWAAAAAATAATTTTTTTTTTAWAWAAA | 1019 |
| QY | 203 | GTGGTAGATTTATTTAGTTCTTAGGTTCGAATCGAAGTTGGTTTAGACATTTAGCCCTTAT | 262 |
| DB | 1018 | TAWATTTTAWAAWTTWATWATWATTTTAAATTTTAAATTTTATATWTTTWWAAAT | 959 |
| QY | 263 | CTTTTTTCTAACCAAAATAAATGTAATCGAAAAACCTTTAGAAAAAAAAGAAATCAAAA | 322 |
| DB | 958 | TTTTTTTWTATATTTTATTTWAAAAAAWATTTWAAATTTWAAAAATAATAAATAAAAAA | 899 |
| QY | 323 | TTGAAAAACATATCGGTCGGAGTCGAGAACCCACACCCACCGTGACCCACAATATAAAA | 382 |
| DB | 898 | TTTTTTTAATTTTAWTTATATAATTWATWAAATTAATAATTTTATTTTAAATTTT | 839 |
| QY | 383 | ATAAGAGTTTGTCCTACAGTAAAGGATCTTTTTTATTCCAATCTCTTTTCCACTCTA | 442 |
| DB | 838 | TTTTTAAATTTAWAAWTTTAAAA---WTTTTTTWATTTTGTWATTTWAAATATTTA | 783 |
| QY | 443 | AAATCTTGGAGATTTGCACCGTTAACTAATTAAGTCGTATATCCACCGTCTCTAAAAAAA | 502 |
| DB | 782 | WTAWATWATATTTTATAGWTATATATATATATATATATATATTAATCTAATTA | 723 |
| QY | 503 | CTTGTGTACCGTGCCTCACAATTTCAACTTTGGCCACCCTG-----AAAGCGGTTATGT | 555 |
| DB | 722 | TWAAATTAATAWATAAAATATATATATATATATTAATCTAATTAATAAAWATATWATTTW | 663 |
| QY | 556 | TTAGGTTAGTGTTGCAACAGTTTGAAGCGCATCACTCAGGAGGCTACTTGGCTTCTGCTTT | 615 |
| DB | 662 | TYWATTTAATTTATATTTWATTTATTTAAAAWATTTATAAAAAATAAAWAAWAAAAA | 603 |
| QY | 616 | TGCGTCTTTTGTTCATTTTTTTCAGTGATTTTCTGTGTAACACGGGTACTTCAAAA--CT | 673 |
| DB | 602 | WATAAAAWTTTWWAAAAAANAANAATWATTTTWTTTTAAAAATTTTAAAAAAATWAT | 543 |
| QY | 674 | TATTATAAATPACATAATTTTATAAGTTTCACCTTCATTATATAATCTCATATATAATATA | 733 |
| DB | 542 | TAATTTWAAATTAATAAATWTTWAAWWTATWAAWATTTTAAAAATTTAAAWAAATTTT | 483 |
| QY | 734 | GGGTTTAGATGCCAATTTTTTAAAAAAGAAATAAAAAATAAATAAGATAAATAATCGAAAA | 793 |
| DB | 482 | AATAATTTAAAWTTWAAATAAATTTAATAAATAAATAAATAATWASSVVAAAAATTTTAA | 423 |
| QY | 794 | AATCAAAATGTAAAAAATTTTGGAGGGGACAAATAAATAATCAAAAGTCTATTATTTAAAT | 853 |

us-09-743-823-21.rst

Page 10

[illegible]

Search completed: February 24, 2004, 11:07:12
Job time : 4772.54 secs

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|-------|-------------|--------|----|---------------------|--------------------|
| 1 | 64.4 | 4.7 | 6106 | 14 | US-10-311-455-1446 | Sequence 1446, Ap |
| 2 | 64.4 | 4.7 | 6106 | 16 | US-10-257-186-114 | Sequence 114, App |
| 3 | 64.2 | 4.7 | 7560 | 14 | US-10-311-455-1196 | Sequence 1196, Ap |
| C 4 | 63.4 | 4.6 | 7195 | 14 | US-10-239-676-30 | Sequence 30, Appl |
| C 5 | 63.4 | 4.6 | 7195 | 14 | US-10-240-453-40 | Sequence 40, Appl |
| C 6 | 63.2 | 4.6 | 318 | 10 | US-09-814-353-6389 | Sequence 6389, Ap |
| C 7 | 63.2 | 4.6 | 318 | 10 | US-09-814-353-12666 | Sequence 12666, Ap |
| C 8 | 63.2 | 4.6 | 403 | 10 | US-09-814-353-17949 | Sequence 17949, A |
| C 9 | 63.2 | 4.6 | 416 | 9 | US-09-960-352-4584 | Sequence 4584, Ap |
| C 10 | 62.2 | 4.6 | 375 | 9 | US-09-960-352-15014 | Sequence 15014, A |
| C 11 | 61.2 | 4.5 | 5511 | 14 | US-10-311-455-1974 | Sequence 1974, Ap |
| C 12 | 61 | 4.5 | 385 | 10 | US-09-814-353-19050 | Sequence 19050, A |
| C 13 | 60.4 | 4.4 | 7631 | 14 | US-10-311-455-833 | Sequence 833, App |
| C 14 | 60.2 | 4.4 | 6487 | 12 | US-10-221-613-227 | Sequence 227, App |
| C 15 | 60 | 4.4 | 6254 | 14 | US-10-311-455-1594 | Sequence 1594, Ap |

Db 857 TTTTATTTTATATTAATTTTATTTTATTTAGATGAATTAAGAAAATTTTAAATTTTA 916
Qy 121 TTTTAAGGCAATTAAGCATGTTT-GATAAAATATATATATGTTTATAAATACCTTTTCAA 179
Db 917 ATTAAATTTGATTTGTAATTTTATTTAGTAAGATAATATATTTATAGATGATTTGTA 976
Qy 180 AGTATAAGTTGATGATGCGGTGGTGTAGATTTATTTAGTTCTAGGTTGCAATCAAGT 239
Db 977 AATATAAATAATTAAGAAAGATTTGTTATTTTATTTTATTTTATTTGTTGTTTGAAGT 1036
Qy 240 TGGTTT 245
Db 1037 GAGTAT 1042

RESULT 2
US-10-257-166-114
; Sequence 114, Application US/10257166
; Publication No. US20040023230A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; TITLE OF INVENTION: Method and Nucleic Acids for Analysing the Methylation of
; TITLE OF INVENTION: Genes Implicated in Pharmacogenomics
; FILE REFERENCE: 5013.1011
; CURRENT APPLICATION NUMBER: US/10/257,166
; CURRENT FILING DATE: 2002-10-07
; PRIOR APPLICATION NUMBER: PCT/EP01/07470
; DE 10032529.7
; DE 10043826.1
; PRIOR FILING DATE: 2001-06-29
; 2000-06-30
; 2000-09-01
; NUMBER OF SEQ ID NOS: 178
; SEQ ID NO 114
; LENGTH: 6106
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-257-166-114

Query Match 4.7%; Score 64.4; DB 16; Length 6106;
Best Local Similarity 56.5%; Pred. No. 0.16;
Matches 139; Conservative 0; Mismatches 106; Indels 1; Gaps 1;

Qy 1 TAAATAATCTATACATTAATAAATTTGATTTTAAATTTTAAATTTTGAATTCATGATTTTAT 60
Db 797 TATTTTAAAGTTTATTTTATTTAGAGTTAATTAATAAATCGCGTAAATAAATTTTATTTT 856
Qy 61 TTTTATTTTACCAGAAATCCGTTAATTTGTTAAATATTACCACTAATTTATATATTTTA 120
Db 857 TTTTATTTTATATATATTTTATTTTATGAGATGAATTAAGAAAATTTTAAATTTTA 916
Qy 121 TTTTAAGGCAATTAAGCATGTTT-GATAAAATATATATATGTTTATAAATACCTTTCAA 179
Db 917 ATTAAATTTGATTTGTAATTTTATTTAGTAAGATAATATATTTATAGATGATTTGTA 976
Qy 180 AGTATAAGTTGATGATGCGGTGGTGTAGATTTATTTAGTTCTAGGTTGCAATCAAGT 239
Db 977 AATATAAATAATTAAGAAAGATTTGTTATTTTATTTTATTTTATTTGTTGTTTGAAGT 1036
Qy 240 TGGTTT 245
Db 1037 GAGTAT 1042

RESULT 3
US-10-311-455-1196
; Sequence 1196, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:

; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determini
; TITLE OF INVENTION: cytosine methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 1196
; LENGTH: 7560
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-1196

Query Match 4.7%; Score 64.2; DB 14; Length 7560;
Best Local Similarity 50.5%; Pred. No. 0.18;
Matches 156; Conservative 0; Mismatches 153; Indels 0; Gaps 0;

Qy 1 TAAATAATCTATACATTAATAAATTTGATTTTAAATTTTAAATTTTGAATTCATGATTTTAT 60
Db 2072 TAAATAAAGTAAGAAATGATTTATTTATTTTATTTTATTTTATTTTATTTTATTTTAT 2131
Qy 61 TTTTATTTTACCAGAAATCCGTTAATTTGTTAAATATTACCACTAATTTATATATTTTA 120
Db 2132 AATTTTATTTGTAAGTTAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 2191
Qy 121 TTTTAAGGCAATTAAGCATGTTTGTATAAATAATATATATTTTGTATAAATACCTTTCAA 180
Db 2192 GTTTTATGATTTAGAAAAGATTTTAAAGATTTTAAATTTGTTTATATAATGTAATAAT 2251
Qy 181 GTATAAGTTGATGATGCGGTGGTGTAGATTTATTTAGTTCTAGGTTGCAATGCAAGT 240
Db 2252 TTTTATTTTATTTTATTTGATATATTTTAAATAAATGAGTATTTATGTTTATTTATTA 2311
Qy 241 GGTTTAGACATTTAGCCTTATTTCTTTTCTAACCAAAATAAATGTAATGGAACCTT 300
Db 2312 GATTTGAATTTAAATAATTTATATTTTAAATAAAGATTTGAAAATAGGAAAAGA 2371
Qy 301 TAGAAAAA 309
Db 2372 AAAGATAAA 2380

RESULT 4
US-10-239-676-30/c
; Sequence 30, Application US/10239676
; Publication No. US20030082609A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; TITLE OF INVENTION: Diagnosis of Diseases Associated with Gene Regulation
; FILE REFERENCE: 5013.1003
; CURRENT APPLICATION NUMBER: US/10/239,676
; CURRENT FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: PCT/EP01/03968
; DE 10019058.8
; DE 10019173.8
; DE 10032529.7
; DE 10043826.1
; PRIOR FILING DATE: 2001-04-06
; 2000-04-06
; 2000-04-07
; 2000-06-30

```
2000-09-01
; NUMBER OF SEQ ID NOS: 228
; SEQ ID NO 30
; LENGTH: 7195
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-239-676-30

Query Match
Best Local Similarity 4.6%; Score 63.4; DB 14; Length 7195;
Matches 157; Conservative 0; Mismatches 156; Indels 0; Gaps 0;

675 ATTATAAATTCACATAATTTTATAGTTTCACCTCTTATATAAATCTCTATATATATATATAG 734
6759 ATTATCCAAAACCAATCATTTTATTAATTTTATTTTAAACACTATATATTTTAAAT 6700
6735 GGTTTAGAAATGCCAATTTTAAACAAAAGAAATAAATAAATAAATAAATAAATAAATAAATAA 794
6699 TATTTAAATAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 6640
6795 ATGAAATGTAAAAATTTGAGGGGACAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 854
6639 AAAACATTTTAAATAATTTTCAACAAAAACAAAAAATTTTAAACAAAAACATCTAAAC 6580
855 CCATTAGATTCCTATTTTCCCTTAGTTTATATGAGTACCCAGTGGGAGATACAGAAA 914
6579 CCAATTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 6520
915 TGTCTAGAAACAGTTGCGATGTAGGGAAATTAATGTAGTAGGGGATAGCAAGACAAAAAT 974
6519 ACTAAGCAATCAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 6460
975 CCAAGCCAAAGCTA 987
6459 ACCCAACAAAATA 6447

RESULT 6
US-09-814-353-6389/c
; Sequence 6389, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-0068
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6389
; LENGTH: 318
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 121, 123, 124, 125, 126, 129, 134, 157, 158, 159, 160, 161,
; LOCATION: 172, 179, 180, 183, 194, 205, 210, 211, 213, 216, 219, 220,
; LOCATION: 225, 226, 227, 235, 241, 243, 248, 250, 267, 268, 272, 277,
; LOCATION: 282, 287, 293, 302
; OTHER INFORMATION: n = A,T,C or G
US-09-814-353-6389

Query Match
Best Local Similarity 4.6%; Score 63.2; DB 10; Length 318;
Matches 98; Conservative 0; Mismatches 75; Indels 0; Gaps 0;
```

666 TTGAAACTTATTATTAAATTTACATAATTTTATAAGTTTCACCTCTCTATATAATTAAGTACTCATAT 725
 200 TTTAAANTTTTTAAAAANAANNNTTTTTTNAAAAAATTTTNNNNNTTTTTAAAAAATAAAAA 141
 726 AATATATAGGGTTTAGAATGCCAATTTTAAABAAAGAAATAAAAAAATTAATAGATATAA 785
 140 AAAAAANTGGGNCNNNTNTTTTTTAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 81
 786 ATCGAAAAAATGAAATGTAAAAAATTTGAGGGGACAAATAAAATATGAAAGT 838
 80 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAGT 28

```

RESULT 7
US-09-814-353-12666/c
; Sequence 12666, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER

```

```

;   OTHER INFORMATION: n = A,T,C or G
US-09-814-353-12666

Query Match      4.6%; Score 63.2; DB 10; Length 318;
Best Local Similarity 56.6%; Pred. No. 0.08;
Matches 98; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      666 TTGAAACTTATTATAAAATTCATAAATTTTATAAGTTTCACCTTTCTATATAATACATCATAT 725

```

RESULT 8

```

US-09-814-353-17949
; Sequence 17949, Application US/09814353
; Publication NO. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,561
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,572
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17949
; LENGTH: 403
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 225
; OTHER INFORMATION: n = A,T,C or G
US-09-814-353-17949

```

RESULT 9
US-95-960-352-4584
Sequence 4584, Application US/09960352
Patent No. US20020137139A1
GENERAL INFORMATION:
APPLICANT: Warren, Wesley C.
APPLICANT: Tao, Nengbing
APPLICANT: Byatt, John C.
APPLICANT: Mathialagan, Nagappan
TITLE OF INVENTION: NUCLEIC ACID AND
TITLE OF INVENTION: MUSCLE AND FAT


```

; Sequence 19050, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19050
; LENGTH: 385
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-814-353-19050

Query Match      4.5%; Score 61; DB 10; Length 385;
Best Local Similarity 59.5%; Pred. No. 0.21; Indels 0; Gaps 0;
Matches 103; Conservative 0; Mismatches 70;

QY 666 TTGAACCTATTATAAATACATAATTTATAGTTTCCTCTTATATACATCATAT 725
Db 268 TTTAAATTTTTTAAAAAATTTTTTTAAAAAATTTTTTGTGTTTTTAAAAA 209
QY 726 AATATATAGGTTTAGAATGCCAATTTTAAAAAGATAAAAAATAGATATAA 785
Db 208 AAAAAATGGGCCATGTTTTTTTTTAAAAAATAAAAAATCAAAAAAAGT 149
QY 786 ATCGAAAAATGAATGTAAAAAATTTGAGGGGACAAATAAATATGAAAGT 838
Db 148 AAAAAAATGAAAAAATCAAAAAAATCAAAAAAATCAAAAAAATCAAAAAA 96

RESULT 13
US-10-311-455-833/c
; Sequence 833, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determ
; TITLE OF INVENTION: cytosine methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 833
; LENGTH: 7631
; TYPE: DNA
; ORGANISM: Artificial Sequence

```

```

; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; US-10-311-455-833

Query Match      4.4%; Score 60.4; DB 14; Length 7631;
Best Local Similarity 61.4%; Pred. No. 0.85;
Matches 97; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

QY 655 AACACGGTACTTGAACTTATTATAAATACATAAATTTTATAGTTTCACCTCTTATAT 714
Db 6669 AAAAACTTTTCGTAAAAATTTAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 6510
QY 715 AATACTCATATAATATAGGTTTGAATGCCAATTTTAAAAAAGATAAAAAAATA 774
Db 6609 ATAACATAATAATTTTAAAAAACAATACTCTCTTTTAAAAAATAAAAAAATA 6550
QY 775 AATAGATAAATCGAAAAATGAAATGTAAAAAATTT 812
Db 6549 AAAATTCCTAACTCTAAAAAACAATAAAAAAATAAAAAAATAAAAAAATA 6512

RESULT 14
US-10-221-613-227
; Sequence 227, Application US/10221613
; Publication No. US20040029123A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with Cell Cycle
; FILE REFERENCE: 5013.1004
; CURRENT APPLICATION NUMBER: US/10/221,613
; CURRENT FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: PCT/EP01/02945
; DE 10013847.00
; DE 10019058.8
; DE 10019173.8
; DE 10032529.7
; DE 10043826.1
; PRIOR FILING DATE: 2001-03-15
; 2000-03-15
; 2000-04-06
; 2000-04-07
; 2000-06-30
; 2000-09-01
; NUMBER OF SEQ ID NOS: 428
; SEQ ID NO 227
; LENGTH: 6487
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; US-10-221-613-227

Query Match      4.4%; Score 60.2; DB 12; Length 6487;
Best Local Similarity 50.7%; Pred. No. 0.87;
Matches 170; Conservative 0; Mismatches 163; Indels 2; Gaps 1;

QY 11 TATACATTAATAAATTTGATTTTAAAAATTTTAGAAATTCATGATTTTATTTTTTACC 70
Db 4081 TAAAGATAATTTTTTTTTTTTAAAAATTTGTAATTTGTAAGTAAGTTTATTTTAA 4140
QY 71 AGAAATCCGTTAATATGTAATAATATTACCACTAATTTATATAATTTTATTAAGCA 130
Db 4141 AGTTAAAGTTTTTTTATTTTAAAAAGTATATATTAGTAAATTTTATTTTATTTGAAG 4200
QY 131 ATTAAGCATGTTTGTAAAAAATATATATTTGTTTATTAATCTTTTCA--AAAGTATAAG 188
Db 4201 TATTTTAAATTTTGGATATATAAGTTAAAAATTTTAAAGTATAAGGAAAGTTTAAA 4260
QY 189 TTGATGATGCGTGGTGGTAGATTATTTTAGTTCTAGTTTCGAATCGAAGTTGGTTTGA 248
Db 4261 ATAATAAGCGTTTATTTTGTGTTATTTTATTTTAAAGAGATTATTTATTTTGTAGGT 4320

```


Qy 249 CATTAGCTTATCTTTTCTAACCAAAATAATGTAATGAAACCTTTAGGAAA 308
Db |||||
Qy 4321 GTTTTAAATAATGGTGTGTTTCTTATATAATAGTAATGCGTGGAGTGATACGG 4380
Db |||||
Qy 309 AAAAGAAATCAAAATGAAACATCATCCCGTGCA 343
Db |||||
Qy 4381 AATAGTATTAGTAAGACGAAATATAATAATAGGA 4415
Db |||||

RESULT 15

US-10-311-455-1594
; Sequence 1594, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIERENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining Cytosine Methylation
; FILE OF INVENTION: cytosine methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311.455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 1594
; LENGTH: 6254
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-1594

Query Match 4.4%; Score 50; DB 14; Length 6254;
Best Local Similarity 53.9%; Pred. No. 0.93;
Matches 166; Conservative 0; Mismatches 140; Indels 2; Gaps 2;
Qy 9 TCTATACATTAAAAAATTGATTTTAAATTTTAGAATTCATGATTTTATTTTITTA 68
Db |||||
Qy 2641 TTTTATATTAAAGATTGTTTAAATATTTAATTAAGAAATTTTITTAAGTA 2700
Db |||||
Qy 69 CCAGAAATCCGTTAATATTGTTAAATATTACCACAACTAATTAATAATTTTATTTAAG 128
Db |||||
Qy 2701 TAAGATAAATTTAAATTTTAAATTTAGTAATAAAGTATAAAATTTTATTAGAATT 2760
Db |||||
Qy 129 CAATTAAGCATGTTGATAAATATATATATTGTTATATAATCTTTTCAAAGCTATAAG 188
Db |||||
Qy 2761 AAGTAATAATTTTIG-TATTATAAATAATAATAATTTTAAATAATATTATTATTAAT 2819
Db |||||
Qy 189 TTGATGATGGCGTGGTGTAGATTATTTTA-GTTCTAGGTTCGAATCGAAGTTGGTTAG 247
Db |||||
Qy 2820 TTGGAGTGAGTATTATTGTTATAATATATTATTTTITTTTGTGTTAGAGATGTTTA 2879
Db |||||
Qy 248 ACATTTAGCCTTATCTTTTCTAACCAAAATAAATGTAATGGAACCTTTAGGAAA 307
Db |||||
Qy 2880 TTAGGTAATAATTTATGATTATCAAAAGTGTTTTAAATGTTAATAGGAGATGCAAGAT 2939
Db |||||
Qy 308 AAAAAGAA 315
Db |||||
Qy 2940 ATAAAGAA 2947
Db |||||

Search completed: February 24, 2004, 13:09:15
Job time : 614.765 secs

BEST AVAILABLE COPY

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 24, 2004, 07:40:35 ; Search time 134.614 Seconds
(without alignments)
5635.487 Million cell updates/sec

Title: US-09-743-823-21

Perfect score: 1367
Sequence: 1 taaataatctatcataaa.....ctacacacactagaagaata 1367

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:*

1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|--------------------|
| C 1 | 59 | 4.3 | 6124 | 4 | US-08-213-419B-3 |
| C 2 | 56.4 | 4.1 | 5332 | 4 | US-09-801-861-3 |
| C 3 | 55 | 4.0 | 2246 | 4 | US-09-363-708-3 |
| C 4 | 55 | 4.0 | 2246 | 4 | US-09-083-587-3 |
| C 5 | 54.8 | 4.0 | 640681 | 4 | US-09-790-988-1 |
| C 6 | 54.2 | 4.0 | 5852 | 1 | US-07-867-106-2 |
| C 7 | 53.6 | 3.9 | 615 | 3 | US-08-998-416-186 |
| C 8 | 53.6 | 3.9 | 636 | 3 | US-08-998-416-1137 |
| C 9 | 53.6 | 3.9 | 837 | 3 | US-08-998-416-288 |
| C 10 | 53.4 | 3.9 | 8537 | 4 | US-10-204-708-41 |
| C 11 | 52.6 | 3.8 | 6124 | 4 | US-08-213-419B-3 |
| C 12 | 52.2 | 3.8 | 827 | 3 | US-08-998-416-535 |
| C 13 | 52.2 | 3.8 | 20674 | 4 | US-09-641-638-651 |
| C 14 | 51.6 | 3.8 | 8306 | 4 | US-10-204-708-49 |
| C 15 | 51.6 | 3.8 | 8920 | 2 | US-08-446-855A-1 |
| C 16 | 51.6 | 3.8 | 8920 | 3 | US-09-150-741-1 |
| C 17 | 51 | 3.7 | 662 | 3 | US-08-998-416-185 |
| C 18 | 51 | 3.7 | 663 | 3 | US-08-998-416-191 |
| C 19 | 51 | 3.7 | 665 | 3 | US-08-998-416-937 |
| C 20 | 51 | 3.7 | 701 | 3 | US-08-998-416-701 |
| C 21 | 51 | 3.7 | 711 | 3 | US-08-998-416-786 |
| C 22 | 51 | 3.7 | 724 | 3 | US-08-998-416-683 |
| C 23 | 51 | 3.7 | 732 | 3 | US-08-998-416-1036 |
| C 24 | 51 | 3.7 | 767 | 3 | US-08-998-416-472 |
| C 25 | 51 | 3.7 | 828 | 3 | US-08-998-416-538 |
| C 26 | 51 | 3.7 | 834 | 3 | US-08-998-416-305 |
| C 27 | 50.8 | 3.7 | 6317 | 4 | US-10-204-708-11 |

| | | | | | | |
|------|------|-----|--------|---|--------------------|--------------------|
| 28 | 50.4 | 3.7 | 6243 | 2 | US-09-056-075-1 | Sequence 1, Appli |
| C 29 | 50.4 | 3.7 | 8537 | 4 | US-10-204-708-42 | Sequence 42, Appli |
| C 30 | 50.4 | 3.7 | 19124 | 2 | US-08-487-836B-13 | Sequence 13, Appli |
| C 31 | 50.4 | 3.7 | 26000 | 4 | US-09-843-376-10 | Sequence 10, Appli |
| C 32 | 50.2 | 3.7 | 8093 | 4 | US-10-204-708-31 | Sequence 31, Appli |
| C 33 | 50 | 3.7 | 6669 | 4 | US-10-204-708-6 | Sequence 6, Appli |
| C 34 | 50 | 3.7 | 7218 | 1 | US-08-232-463-14 | Sequence 14, Appli |
| C 35 | 49.8 | 3.6 | 640681 | 4 | US-09-790-988-1 | Sequence 1, Appli |
| C 36 | 49 | 3.6 | 194 | 4 | US-09-621-976-9596 | Sequence 9596, Ap |
| C 37 | 49 | 3.6 | 4673 | 1 | US-07-638-431-1 | Sequence 1, Appli |
| C 38 | 49 | 3.6 | 4673 | 5 | PCT-US92-00018-1 | Sequence 1, Appli |
| C 39 | 49 | 3.6 | 6866 | 4 | US-10-204-708-20 | Sequence 20, Appli |
| C 40 | 49 | 3.6 | 7304 | 4 | US-10-204-708-44 | Sequence 44, Appli |
| C 41 | 49 | 3.6 | 19233 | 4 | US-10-204-708-46 | Sequence 46, Appli |
| C 42 | 48.8 | 3.6 | 1117 | 3 | US-09-247-373B-33 | Sequence 33, Appli |
| C 43 | 48.8 | 3.6 | 1361 | 4 | US-09-489-847-64 | Sequence 64, Appli |
| C 44 | 48.8 | 3.6 | 6070 | 4 | US-10-204-708-9 | Sequence 9, Appli |
| C 45 | 48.8 | 3.6 | 11015 | 4 | US-10-204-708-55 | Sequence 55, Appli |

ALIGNMENTS

RESULT 1

US-08-213-419B-3/c

Sequence 3, Application US/08213419B

Patent No. 6333406

GENERAL INFORMATION:

APPLICANT: Inselburg, J. et al.

TITLE OF INVENTION: GENE ENCODING PROTEIN ANTIGENS OF PLASMODIUM FALCIPARUM

TITLE OF INVENTION: AND USES THEREFOR

FILE REFERENCE: J11-002CNCIP

CURRENT APPLICATION NUMBER: US/08/213,419B

CURRENT FILING DATE: 1994-03-14

PRIOR APPLICATION NUMBER: US 07/870,506

PRIOR FILING DATE: 1992-04-17

NUMBER OF SEQ ID NOS: 20

SOFTWARE: Patent In Ver. 2.0

SEQ ID NO 3

LENGTH: 6124

TYPE: DNA

ORGANISM: Plasmodium falciparum

FEATURE:

NAME/KEY: CDS

LOCATION: (2407)..(2439)

NAME/KEY: CDS

LOCATION: (2598)..(3404)

NAME/KEY: CDS

LOCATION: (3580)..(3720)

NAME/KEY: CDS

LOCATION: (3850)..(5835)

US-08-213-419B-3

Query Match 4.3%; Score 59; DB 4; Length 6124;

Best Local Similarity 56.4%; Pred. No. 0.0043;

Matches 110; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

| | | | |
|----|-----|--|-----|
| QY | 675 | ATTATAAATACATATTTTATAAGTTTCACCTCTTATATACTATCATATATATATAG | 734 |
| Db | 782 | ATATTAAATAAAGAATTAACATATAGTTTAACTAAGACTAGTTTAAATTCATAT | 723 |
| QY | 735 | GGTTTGAATGCCAATTTTAAAAAAGAAATAAATAAATAGAAATCGAAAAA | 794 |
| Db | 722 | TTTAAATATTTAAATGAAAAAATAAATAAATAAATAAATAAATAAATAAATA | 663 |
| QY | 795 | ATGAATGTAAAAAATTTTCAGGGGCAATAAATAATGAAAGTCTATTATTAAATTTT | 854 |
| Db | 662 | ATAATAAGAAACAAACTTGATATAAATTTTCAATATGATTATTATTATATATT | 603 |
| QY | 855 | CCATTAGAATTCAT | 869 |
| Db | 602 | TCATAAAATTTTAT | 588 |

RESULT 2
US-09-801-861-3/c
; Sequence 3, Application US/09801861
; Patent No. 6492154
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001098
; CURRENT APPLICATION NUMBER: US/09/801,861
; CURRENT FILING DATE: 2001-03-09
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 5332
; TYPE: DNA
; ORGANISM: Human
US-09-801-861-3

Query Match 4.1%; Score 56.4; DB 4; Length 5332;
Best Local Similarity 48.4%; Pred. No. 0.02; Indels 0; Gaps 0;
Matches 156; Conservative 0; Mismatches 166; Indels 0; Gaps 0;

QY 1 TAAATAATCTATACATTAATAAATTTGATTTTAAATTTTGAATAATTCATGATTTTATT 60
Db 31641 TAAAAAATATATATAATATATAAATATATAAATATATAAATATATAAATATATAAAT 31582

QY 61 TTTTITACCAGAAATCCGTTAATTTGTTAAATATTTACCAACTAAATTTTAAATTTTA 120
Db 31581 ATATAAATATATAATATATAAATATATAAATATATAAATATATAAATATATAAAT 31522

QY 121 TTTTAAAGGCAATTAAGCATGTTTGATAAATATATATATTTGTTATAAATACCTTTTCAAAA 180
Db 31521 ATATAAATATATAATATATAAATATATAAATATATAAATATATAAATATATAAAT 31462

QY 181 GTATAAAGTTGATGGCGTGGTGGTAGATATTTTGTAGTTTCTAGTTCGAATGCAAGTT 240
Db 31461 ATAAAAATATAAATATAAATAAATATAAATAAATATAAATAAATATAAATAAATAAATA 31402

QY 241 GGTTTAGACATTTAGCTTATTTCTTTCTACCAAAATAAATGTAATGGAACCTT 300
Db 31401 TATATAAATATATAAATATATAAATATATAAATATATAAATATATAAATATATAAATA 31342

QY 301 TAGGAAAAAAGAAATCAAAA 322
Db 31341 TATAAATAAATATATAAATAA 31320

RESULT 3
US-09-363-708-3
; Sequence 3, Application US/09363708
; Patent No. 6399747
; GENERAL INFORMATION:
; APPLICANT: Schmandt, et al.
; TITLE OF INVENTION: NOVEL SHC BINDING PROTEIN
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive/6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/363,708
; FILING DATE:

CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Clough, David W.
REGISTRATION NUMBER: 36,107
REFERENCE/DOCKET NUMBER: 01017/34451
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-6300
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2246 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
DESCRIPTION: /desc = "mouse PAL CDNA"
US-09-363-708-3

Query Match 4.0%; Score 55; DB 4; Length 2246;
Best Local Similarity 57.1%; Pred. No. 0.022;
Matches 100; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 562 GTACTTCAAACTTATATAAATTACATAATTTTATAAGTTTCACTTCTTATATAACTC 721
Db 2049 GTATTTGAACATGCTCTTAAGTATGCTGCTTATATACTTTGCTTCATTTGCTTCAGGCT 2108

QY 722 ATATAATATATAGGTTTGAAGTCCCAATTTTAAAAAAGAAATAAAAAATATAAGAA 781
Db 2109 GTGTATTATATAAGTCTACTTGACCAAAAAAATAAAAAAATAAAAAA 2168

QY 782 TAAATCCAAAAATGAATGTAATAAATTTGAGGGGACAAATAAATAATGAAA 836
Db 2169 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 2223

RESULT 4
US-09-083-587-3
; Sequence 3, Application US/09083587
; Patent No. 6492138
; GENERAL INFORMATION:
; APPLICANT: Schmandt, et al.
; TITLE OF INVENTION: NOVEL SHC BINDING PROTEIN
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive/6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/083,587
; FILING DATE:

CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Clough, David W.
REGISTRATION NUMBER: 36,107
REFERENCE/DOCKET NUMBER: 01017/34451
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-6300
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2246 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA

US-09-083-587-3

DESCRIPTION: /desc = "mouse PAL cDNA"

Query Match 4.0%; Score 55; DB 4; Length 2246;
Best Local Similarity 57.1%; Pred. No. 0.022;
Matches 100; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 662 GTACTGAACTTATATATAATACATAATTTTATAAGTTTCACTTCTTATATAATPACTC 721
Db 2049 GTATTTTGAACATGCTCTTAAGTATGCTCTATATACCTTTGCTTCATTTGCTTCATGGCT 2108

QY 722 ATATATATATAGGTTTGAATGCCAATTTTAAAGAAAGATAAAATAAATAAGAA 781
Db 2109 GTGTATTTATAAAGTGTCTTGACCAAAAAAATAAATAAATAAATAAATAAATAA 2168

QY 782 TAAATCGAAAAAATGAAATCTATAAATTTTGAAGGGGACAAATAAATAATATGAAA 836
Db 2169 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2223

RESULT 5

US-09-790-988-1/c

Sequence 1, Application US/09790988

Patent No. 6632935

GENERAL INFORMATION:

APPLICANT: SHIGENOBU, SHUJI

APPLICANT: WATANABE, HIDEMI

APPLICANT: HATTORI, MASAHIRA

APPLICANT: SAKAKI, YOSHIYUKI

TITLE OF INVENTION: GENOME DNA OF BACTERIAL SYMBIONT OF APHIDS

FILE REFERENCE: 081356/0159

CURRENT APPLICATION NUMBER: US/09/790,988

PRIOR FILING DATE: 2001-02-23

PRIOR FILING DATE: 2000-04-07

NUMBER OF SEQ ID NOS: 7

SOFTWARE: Patent In Ver. 2.1

SEQ ID NO 1

LENGTH: 640681

TYPE: DNA

ORGANISM: Buchnera sp.

US-09-790-988-1

Query Match 4.0%; Score 54.8; DB 4; Length 640681;
Best Local Similarity 42.9%; Pred. No. 0.061;
Matches 381; Conservative 0; Mismatches 502; Indels 5; Gaps 2;

QY 1 TAAATAATCTATACATAAATAATTTGATTTTAAATTTTGAATTTAGAAATTCATGATTTTAT 60
Db 353537 TTAAGTAATATCAATTTTAAAGTAATATTTTAAATAATTTATATGATTTTCTAATAATAGATA 353478

QY 61 TTTTTCACAGAAATCCGTTAATATTTGTTAAATAATTTACCACTA---ATTTTATAAAT 117
Db 353477 AATTTTATATAAATCTTTTATTTTATTTTAAACTAACTAACTAAGTATGATTAAGA 353418

QY 118 TTAATTTAAGGCAATTAAGCATGTTTGATAAATAATATATATTTGTTATAAATACTTTTCA 177
Db 353417 AATCAATTAAGTAAATGATTTATTTTAAATAATTAATATAAATAAATAAATAAATAA 353358

QY 178 AAGATATAAGTTTCATGATCGCGTGGTGTAGATTTATTTAGTTCTAGGTTTCAATGCAA 237
Db 353357 GAATATTTTTCGATGATCATATATTTTAAAGGATTTCTTTTAAATAATATGATA 353298

QY 238 GTTCGTTAGACATTTAGCCCTTATCTTTTCTCAACCAAAATAAATGTAATCGAAAC 297
Db 353297 AAGATTTTAAATAAATACTTTTATTTATTTATTTTATTTATTTTAAATTTATCTT 353238

QY 298 CTTTAGGAAAAAAGAAATCAAAATTTGAAACATCTCCGGTGGAGTCGAGAGCCAC 357
Db 353237 TTTCTGGTATTTATTTCTAAATGATAAACAACAAACGCTAATTTATTTATATAGATAATATAG 353178

QY 358 ACCCAGGTGACCCCAACAATATTAATAAAGAGTTTGTCTACAGTAATCGGATCTTTT 417

Db 353177 AAAAAATGATAATCTCTGTTTTTATAAGATAATATCTTTTATTTTATTTTATTTTATAAATTT 353118

QY 418 TTATTCATATCTTTTCCACTTTTAAATCTTTGGAGATTTTCACCGTTTAACTAATTAAGT 477
Db 353117 TTATTCAGATTTAAGTATTTGTTTAAATAAATATGATAAATAAATTTTAAATAATTTATTT 353058

QY 478 GTTATATCCAAAGGTCCTTAAATAAATCTTTGTACCGTCCCTCACATTTCAACTTTGCGCA 537
Db 353057 TATAAATATCTTTCACAGAAAAATTAATAAGAAAGATATCTTTTAAACATAAATAATCT 352998

QY 538 CCTCGAAAGCCGTTATGTTTAGTTAGTTTGAACAGTTTGAAGCGCATCACTCAGGAG 597
Db 352997 TAAITGAATAGTATCACTTATAGTGCATATTAACAGATTTTAAACATATTTACTACTAT 352938

QY 598 GCTACTTCGCTTCGTTTTCGCTTTTGTTCATTTTTCACGATTTTTCGTTGTTGAAC 657
Db 352937 ATTTTAAATAAATGTTTAAATTTTCTTGAATAATTTAGTATTTATTTTATTTTATGAA 352878

QY 658 AGCGTACTTGAACCTTATTTATATAAATTTTATAAGTTTTCCTTCTTATATAAT 717
Db 352877 TAGGCATATTTTATATATATAAGTATGTTTAAAGATACTATTGATTTAGTATAGATT 352818

QY 718 ACTCATATAATATAGGTTTGAATCCCAATTTTAAATAAATAAATAAATAAATAAATAA 777
Db 352817 TTTATAATATAATTAAGTCTATATAAATTTCTGTTAATTTGATAATTTTATAACA--TATC 352760

QY 778 AGAATAAATCGAAAAATGAATGTAAATAATTTGAGGGGACAAATAAATAATGAAG 837
Db 352759 CCAGTATTTTCAATAATTTTGTGTAACTAACGTTGAACAAAGATAAATAAATAAATAA 352700

QY 838 TCTATTATTTAAATTTCCATTAGAAATCTATTTTCCCTTAGTTAATAT 885
Db 352699 TCAATACTATTTTTCATATAATATTTTATTTTATTTTAAAT 352652

RESULT 6

US-07-867-106-2/c

Sequence 2, Application US/07867106

Patent No. 5389526

GENERAL INFORMATION:

APPLICANT: Slade, Martin B

APPLICANT: Chang, Andy C M

APPLICANT: Williams, Keith L

TITLE OF INVENTION: Improved Plasmid Vectors for Cellular

TITLE OF INVENTION: Slime Moulds of the Genus Dictyostelium

NUMBER OF SEQUENCES: 13

CORRESPONDENCE ADDRESS:

ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5389526ris

STREET: One Liberty Place 46th Floor

CITY: Philadelphia

STATE: PA

COUNTRY: USA

ZIP: 19103

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/867,106

FILING DATE: 19920625

PRIOR APPLICATION DATA:

APPLICATION NUMBER: AU PJ 7187

APPLICATION NUMBER: PCT/AU90/00530

FILING DATE: 02-NOV-1989

ATTORNEY/AGENT INFORMATION:

NAME: Feeney, Joanne Longo

REGISTRATION NUMBER: 35,134

REFERENCE/DOCKET NUMBER: RICE-0002

TELECOMMUNICATION INFORMATION:

TELEPHONE: 215-568-3100

TELEFAX: 215-568-3439

INFORMATION FOR SEQ ID NO: 2:

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 5852 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2378..5038
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2378..5038
; US-07-867-106-2

Query Match
Best Local Similarity 4.0%; Score 54.2; DB 1; Length 5852;
Matches 112; Conservative 0; Mismatches 78; Indels 1; Gaps 1;

QY 683 TTACATAATTTTATAGTTTCACCTTCTTATA-TAATACTCATATAATATATAGGTTTATAG 741
Db 5713 TTACTTATATTTCTATTTTATTATATAAATTAATAATTTTAAATAAATAAATAA 5654

QY 742 ATGCCAATTTTAAAAAAGAAATAAAAAATAAATAAGATAAATAATCGAAAAAATGAAT 801
Db 5653 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 5594

QY 802 GTAAAAAATTCAGGGGGGACAAATAAATAATGCAAGTCTATTATTAAATTTTCCATTAG 861
Db 5593 ATGTATTAATTTGCAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAA 5534

QY 862 AATTCATTTT 872
Db 5533 AATTTTATTTT 5523

```

```

RESULT 7
US-08-998-416-186
; Sequence 186, Application US/08998416
; Patent No. 6239264
; GENERAL INFORMATION:
; APPLICANT: Philippsen, Peter
; APPLICANT: Pohlmann, Rainer
; APPLICANT: Steiner, Sabine
; APPLICANT: Mohr, Christine
; APPLICANT: Wendland, Jurgen
; APPLICANT: Knechtle, Philipp
; APPLICANT: Rebischung, Corinne
; TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPHII
; NUMBER OF SEQUENCES: 1152
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6239264artis Corporation
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: No. 6239264th Carolina
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/998,416
; FILING DATE: 24-DEC-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION NUMBER: CH 0016/97
; FILING DATE: 31-DEC-1996
; APPLICATION DATA:
; NAME: Meigs, J. Timothy
; TELEPHONE: 919-541-8587
; TELEFAX: 919-541-8689
; REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976

```

```

; REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8587
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 186:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 615 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: PAG1074RP
; US-08-998-416-186

```

```

Query Match
Best Local Similarity 3.9%; Score 53.6; DB 3; Length 615;
Matches 107; Conservative 0; Mismatches 69; Indels 2; Gaps 1;

QY 5 ATAATCTATACATTAATAAATTTGATTTTAAATTTTAGAAATTCATGATTTTATTTT 64
Db 26 ATAAAGATTAAATAAATTTTATTATAATAATTTTAAAGTATTAAATTTTAAACTATT 85

QY 65 TTACACGAAATCCGTTAAT--ATTGTTAAATATTACCACTAATTTATAAATTTTAT 122
Db 86 ATTATCATTTTAAATAAATTTTATTATTGATTTTAACTATTATTATAATATTATA 145

QY 123 TTAAGGCAATTAAGCATGTTTGATAAATAATATATATTTGTTATAAATCTTTTCAAAA 180
Db 146 TAATTTACITTAATTCATCATTTATTATATTATATTATAAATAATATTATAATA 203

```

```

RESULT 8
US-08-998-416-1137
; Sequence 1137, Application US/08998416
; Patent No. 6239264
; GENERAL INFORMATION:
; APPLICANT: Philippsen, Peter
; APPLICANT: Pohlmann, Rainer
; APPLICANT: Steiner, Sabine
; APPLICANT: Mohr, Christine
; APPLICANT: Wendland, Jurgen
; APPLICANT: Knechtle, Philipp
; APPLICANT: Rebischung, Corinne
; TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPHII
; NUMBER OF SEQUENCES: 1152
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6239264artis Corporation
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: No. 6239264th Carolina
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/998,416
; FILING DATE: 24-DEC-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION NUMBER: CH 0016/97
; FILING DATE: 31-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8587
; TELEFAX: 919-541-8689

```

```
; INFORMATION FOR SEQ ID NO: 1137:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 636 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: PAG1692RP
US-08-998-416-1137

Query Match          3.9%; Score 53.6; DB 3; Length 636;
Best Local Similarity 60.1%; Pred. No. 0.035;
Matches 107; Conservative 0; Mismatches 69; Indels 2; Gaps 1;

QY 5 ATATCTATACATTAATAAATTTGATTTTAAATTTTAGAAATTCATGATTTTATTTT 64
   |||||
Db 26 ATAAAGATTATATATAACTTTTATTTATTAATTTTAAAGTATTAAATTTTAACTATT 85
   |||||

QY 65 TTACACAGAAATCGTTAAT--ATTGTTAAATATACCAACTAATTTTATAAATTTTATT 122
   |||||
Db 86 ATTATCATTTTAAATAAATTTTATTTGATTATTATTAATTAATTAATTTATTA 145
   |||||

QY 123 TTAAGGCAATTAAGCATGTTTGATTAATAATATATATATTTGTTATAAATACTTTCAAAA 180
   |||||
Db 146 TAATTTACTTAATCATCATTTATTAATTTTATTAATTAATTAATAAATTAATTAATA 203
   |||||

RESULT 9
US-08-998-416-288
; Sequence 288, Application US/08998416
; Patent No. 6239264
; GENERAL INFORMATION:
; APPLICANT: Philippsen, Peter
; APPLICANT: Pohlmann, Rainer
; APPLICANT: Stehner, Sabine
; APPLICANT: Mohr, Christine
; APPLICANT: Wendland, Jurgen
; APPLICANT: Knechtie, Philipp
; APPLICANT: Rebschuing, Corinne
; TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPHII
; NUMBER OF INVENTION: AND USES THEREOF
; NUMBER OF SEQUENCES: 1152
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6239264artis Corporation
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: No. 6239264th Carolina
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/998,416
; FILING DATE: 24-DEC-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: CH 0016/97
; FILING DATE: 31-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: PP/5-30306/A/CGC1976
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8587
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 288:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 837 base pairs
; TYPE: nucleic acid
```

```
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: PAG1241RP
US-08-998-416-288

Query Match          3.9%; Score 53.6; DB 3; Length 837;
Best Local Similarity 60.1%; Pred. No. 0.036;
Matches 107; Conservative 0; Mismatches 69; Indels 2; Gaps 1;

QY 5 ATAACTATACATTAATAAATTTGATTTTAAATTTTAGAAATTCATGATTTTATTTT 64
   |||||
Db 26 ATAAAGATTATATATAACTTTTATTTATTAATTTTAAAGTATTAAATTTTAACTATT 85
   |||||

QY 65 TTACACAGAAATCCGTTAAT--ATTGTTAAATATACCAACTAATTTTATAAATTTTATT 122
   |||||
Db 86 ATTATCATTTTAAATAAATTTTATTTGATTATTATTAATTAATTAATTTATTA 145
   |||||

QY 123 TTAAGGCAATTAAGCATGTTTGATTAATAATATATATATTTGTTATAAATACTTTCAAAA 180
   |||||
Db 146 TAATTTACTTAATCATCATTTATTAATTTTATTAATTAATTAATAAATTAATTAATA 203
   |||||

RESULT 10
US-10-204-708-41/c
; Sequence 41, Application US/10204708
; Patent No. 6677731
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
; TITLE OF INVENTION: By Assessing DNA Methylation
; FILE REFERENCE: 5013.1012
; CURRENT APPLICATION NUMBER: US/10/204,708
; CURRENT FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: PCT/EP01/03971
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 41
; LENGTH: 8537
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-41

Query Match          3.9%; Score 53.4; DB 4; Length 8537;
Best Local Similarity 53.6%; Pred. No. 0.058;
Matches 111; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

QY 667 TGAACCTTATTAATAATTAACATAATTTTATAAGTTTCACTTCTTATATAAATCACTATA 726
   |||||
Db 962 TAAAAATATTTATAAATATATTAATTAATAAATTTTAAAAATATAATACTCACACTATA 903
   |||||

QY 727 ATATATAGGTTTGAAGATGCCAATTTTAAAAAAGATAAAAAATTAATAGATAATAA 786
   |||||
Db 902 ATCTAATACATTAATAAATCACTAATAATAAATTAATAAATTAATAAATTAATAAATACTA 843
   |||||

QY 787 TCGAAAAAATGAATGTAAAAATTTGAGGGGCAATAAATAATTAAGAGCTATTATT 846
   |||||
Db 842 ATCTAAACAATATAATAAATTAATAAATTAATAAATTAATAAATTAATAAATTAATAAATTA 783
   |||||

QY 847 TAAATTTCCATTAGAAATTCATTTC 873
```

| | | | |
|---|------|--|------|
| | 782 | TATATTTTTATTTAAATAAATCTC | 756 |
| Db | | | |
| RESULT 11 | | | |
| US-08-213-419B-3 | | | |
| ; Sequence 3, Application US/08213419B | | | |
| ; Patent No. 6333406 | | | |
| ; GENERAL INFORMATION: | | | |
| ; APPLICANT: Inselburg, J. et al. | | | |
| ; TITLE OF INVENTION: GENE ENCODING PROTEIN ANTIGENS OF PLASMODIUM FALCIPARUM | | | |
| ; FILE REFERENCE: JII-002CNCP | | | |
| ; CURRENT APPLICATION NUMBER: US/08/213,419B | | | |
| ; CURRENT FILING DATE: 1994-03-14 | | | |
| ; PRIOR APPLICATION NUMBER: US 07/870,506 | | | |
| ; PRIOR FILING DATE: 1992-04-17 | | | |
| ; NUMBER OF SEQ ID NOS: 20 | | | |
| ; SOFTWARE: PatentIn Ver. 2.0 | | | |
| ; SEQ ID NO 3 | | | |
| ; LENGTH: 6124 | | | |
| ; TYPE: DNA | | | |
| ; ORGANISM: Plasmodium falciparum | | | |
| ; FEATURE: | | | |
| ; NAME/KEY: CDS | | | |
| ; LOCATION: (2407)..(2439) | | | |
| ; NAME/KEY: CDS | | | |
| ; LOCATION: (2598)..(3404) | | | |
| ; NAME/KEY: CDS | | | |
| ; LOCATION: (3580)..(3720) | | | |
| ; NAME/KEY: CDS | | | |
| ; LOCATION: (3850)..(5835) | | | |
| US-08-213-419B-3 | | | |
| Query Match 3.8%; Score 52.6; DB 4; Length 6124; | | | |
| Best Local Similarity 52.0%; Pred. No. 0.079; | | | |
| Matches 118; Conservative 0; Mismatches 109; Indels 0; Gaps 0; | | | |
| QY | 710 | TATATAACTCATATAATATAGGGTTTAGAATGCCAATTTTTAAAAAGAATAAAA | 769 |
| Db | 2089 | TAAAAAGAGTCATATCTAGATTATTCTCTATGAGAACTACAAAAA | 2148 |
| QY | 770 | AAATAAATGAATAAATCGAAAAAAGAAATGTAATAAATTTGAGGGGCACAATAAAA | 829 |
| Db | 2149 | AAATTAATAAATAAATAAATAAATAAATTTAAATAATAATAATAATTTATATATATA | 2208 |
| QY | 830 | TATGAAGTCATATTATTTAAATTTCCCAATAGAAATCTATTTCCCTAGTTTAATATGAGC | 889 |
| Db | 2209 | TTTATATATATATATATATAATAAATTTTTTAATTTAAATAATAATAATATAGATGTCCAAA | 2268 |
| QY | 890 | TAGCCAGTTGGGAGATACACGAAAAATGTCAGAACACAGTTCAGTGA | 936 |
| Db | 2269 | AAAAAATAAATAAATAAATAATATATATATATATAAATAACATATA | 2315 |
| RESULT 12 | | | |
| US-08-998-416-535 | | | |
| ; Sequence 535, Application US/08998416 | | | |
| ; Patent No. 6239264 | | | |
| ; GENERAL INFORMATION: | | | |
| ; APPLICANT: Philippsen, Peter | | | |
| ; APPLICANT: Pohlmann, Rainer | | | |
| ; APPLICANT: Steiner, Sabine | | | |
| ; APPLICANT: Mohr, Christine | | | |
| ; APPLICANT: Wendland, Jurgen | | | |
| ; APPLICANT: Knechtle, Philipp | | | |
| ; APPLICANT: Reibischung, Corinne | | | |
| ; TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPHII | | | |
| ; TITLE OF INVENTION: AND USES THEREOF | | | |
| ; NUMBER OF SEQUENCES: 1152 | | | |
| ; CORRESPONDENCE ADDRESS: | | | |
| ; ADDRESSEE: No. 6239264artis Corporation | | | |
| ; STREET: 3054 Cornwallis Road | | | |


```

; PRIOR APPLICATION NUMBER: US 09/502,330
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: US 60/133,200
; PRIOR FILING DATE: 1999-05-07
; PRIOR APPLICATION NUMBER: US 09/275,267
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: US 60/119,917
; PRIOR FILING DATE: 1999-02-12
; NUMBER OF SEQ ID NOS: 1304
; SOFTWARE: Patent.pm
; SEQ ID NO 651
; LENGTH: 20674
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1123..3123
; OTHER INFORMATION: 5'regulatory region
; NAME/KEY: exon
; LOCATION: 3124..3297
; OTHER INFORMATION: exon 1
; NAME/KEY: exon
; LOCATION: 3298..4072
; OTHER INFORMATION: exon 2
; NAME/KEY: exon
; LOCATION: 4073..5633
; OTHER INFORMATION: exon 3
; NAME/KEY: exon
; LOCATION: 5634..5880
; OTHER INFORMATION: exon 4
; NAME/KEY: exon
; LOCATION: 5881..6099
; OTHER INFORMATION: exon 5
; NAME/KEY: exon
; LOCATION: 6100..6509
; OTHER INFORMATION: exon 6
; NAME/KEY: exon
; LOCATION: 6510..7522
; OTHER INFORMATION: exon 7
; NAME/KEY: exon
; LOCATION: 7523..8854
; OTHER INFORMATION: exon 8
; NAME/KEY: exon
; LOCATION: 8855..12340
; OTHER INFORMATION: exon 9
; NAME/KEY: exon
; LOCATION: 12341..13023
; OTHER INFORMATION: exon 10
; NAME/KEY: exon
; LOCATION: 13024..13429
; OTHER INFORMATION: exon 11
; NAME/KEY: exon
; LOCATION: 13430..16667
; OTHER INFORMATION: exon 12
; NAME/KEY: exon
; LOCATION: 16668..16945
; OTHER INFORMATION: exon 13
; NAME/KEY: exon
; LOCATION: 16946..17554
; OTHER INFORMATION: exon 14
; NAME/KEY: misc_feature
; LOCATION: 17555..20674
; OTHER INFORMATION: 3'regulatory region
; NAME/KEY: allele
; LOCATION: 1128
; OTHER INFORMATION: 10-508-191 : polymorphic base C or T
; NAME/KEY: allele
; LOCATION: 1182
; OTHER INFORMATION: 10-508-245 : polymorphic base C or T
; NAME/KEY: allele
; LOCATION: 1559
; OTHER INFORMATION: 10-509-284 : polymorphic base C or T
; NAME/KEY: allele
; LOCATION: 1570
; OTHER INFORMATION: 10-509-295 : deletion of C
; NAME/KEY: allele
; LOCATION: 1827
; OTHER INFORMATION: 10-510-173 : variable motif ATTGA or TTTTTT
; NAME/KEY: allele
; LOCATION: 2048
; OTHER INFORMATION: 10-511-62 : polymorphic base C or T
; NAME/KEY: allele
; LOCATION: 2323
; OTHER INFORMATION: 10-511-337 : insertion of T
; NAME/KEY: allele
; LOCATION: 2341
; OTHER INFORMATION: 10-512-36 : polymorphic base G or C
; NAME/KEY: allele
; LOCATION: 2623
; OTHER INFORMATION: 10-512-318 : polymorphic base A or G
; NAME/KEY: allele
; LOCATION: 2832
; OTHER INFORMATION: 10-513-250 : polymorphic base A or G
; NAME/KEY: allele
; LOCATION: 2844
; OTHER INFORMATION: 10-513-262 : polymorphic base C or T
; NAME/KEY: allele
; LOCATION: 2934
; OTHER INFORMATION: 10-513-352 : polymorphic base A or G
; NAME/KEY: allele
; LOCATION: 2947
; OTHER INFORMATION: 10-513-365 : polymorphic base A or G
; NAME/KEY: allele
; LOCATION: 3802
; OTHER INFORMATION: 12-206-81 : polymorphic base A or G
; NAME/KEY: allele
; LOCATION: 4062
; OTHER INFORMATION: 10-343-231 : deletion of C
; NAME/KEY: allele
; LOCATION: 4088
; OTHER INFORMATION: 12-206-366 : polymorphic base C or T
; NAME/KEY: allele
; LOCATION: 4109
; OTHER INFORMATION: 10-343-278 : polymorphic base C or T
; NAME/KEY: allele
; LOCATION: 4170
; OTHER INFORMATION: 10-343-339 : polymorphic base G or T
; NAME/KEY: allele
; LOCATION: 5903
; OTHER INFORMATION: 10-346-23 : polymorphic base A or G
; NAME/KEY: allele
; LOCATION: 6019
; OTHER INFORMATION: 10-346-141 : polymorphic base A or G
; NAME/KEY: allele
; LOCATION: 6141
; OTHER INFORMATION: 10-346-263 : polymorphic base G or C
; NAME/KEY: allele
; LOCATION: 6183
; OTHER INFORMATION: 10-346-305 : polymorphic base C or T
; NAME/KEY: allele
; LOCATION: 6338
; OTHER INFORMATION: 10-347-74 : polymorphic base A or G
; NAME/KEY: allele
; LOCATION: 6375
; OTHER INFORMATION: 10-347-111 : polymorphic base G or C
; NAME/KEY: allele
; LOCATION: 6429
; OTHER INFORMATION: 10-347-165 : polymorphic base C or T
; NAME/KEY: allele
; LOCATION: 6467
; OTHER INFORMATION: 10-347-203 : polymorphic base A or G
; NAME/KEY: allele
; LOCATION: 6484
; OTHER INFORMATION: 10-347-220 : polymorphic base A or G
; NAME/KEY: allele
; LOCATION: 6534
; OTHER INFORMATION: 10-347-220 : polymorphic base A or G

```

OTHER INFORMATION: 10-347-271 : polymorphic base A or T
 NAME/KEY: allele
 LOCATION: 6611
 OTHER INFORMATION: 10-347-348 : polymorphic base A or G
 NAME/KEY: allele
 LOCATION: 7668
 OTHER INFORMATION: 10-348-391 : polymorphic base A or G
 NAME/KEY: allele
 LOCATION: 8608
 OTHER INFORMATION: 10-349-47 : polymorphic base C or T
 NAME/KEY: allele
 LOCATION: 8658
 OTHER INFORMATION: 10-349-97 : polymorphic base A or G
 NAME/KEY: allele
 LOCATION: 8703
 OTHER INFORMATION: 10-349-142 : polymorphic base G or C
 NAME/KEY: allele
 LOCATION: 8777
 OTHER INFORMATION: 10-349-216 : deletion of CTG
 NAME/KEY: allele
 LOCATION: 8785
 OTHER INFORMATION: 10-349-224 : polymorphic base G or T
 NAME/KEY: allele
 LOCATION: 8926
 OTHER INFORMATION: 10-349-368 : polymorphic base C or T
 NAME/KEY: allele
 LOCATION: 12171
 OTHER INFORMATION: 10-350-72 : polymorphic base C or T
 NAME/KEY: allele
 LOCATION: 12429
 OTHER INFORMATION: 10-350-332 : polymorphic base C or T
 NAME/KEY: allele
 LOCATION: 13341
 OTHER INFORMATION: 10-507-170 : polymorphic base A or G
 NAME/KEY: allele
 LOCATION: 13492
 OTHER INFORMATION: 10-507-321 : polymorphic base A or C
 NAME/KEY: allele
 LOCATION: 13524
 OTHER INFORMATION: 10-507-353 : polymorphic base C or T
 NAME/KEY: allele
 LOCATION: 13535

Query Match 3.8%; Score 52.2; DB 4; Length 20674;
 Best Local Similarity 47.2%; Pred. No. 0.12;
 Matches 159; Conservative 0; Mismatches 178; Indels 0; Gaps 0;

Qy 16 ATTAAAAATTTGATTTTAAATTTTAGAAATTCATGATTTTATTTTATTTTACCAGAAA 75
 Db 11200 AATTAAATTTAAATTTAAATTTTAAATTAACATTTAAATTTTAAATTTTAAATTTAAAT 11259
 Qy 76 TCCGTTAATATTGTTAAATATTACCACTAATTTTATAAATTTTATTTTAAAGGCAATTA 135
 Db 11260 TAATTTAATTTAATTTAATTTTAACTTAATTTTAAATTTAATTTAATTTTAACTTAATTTAA 11319
 Qy 136 GCATGTTGCAATAATATATATGTTATAATATCTTTTCAAAAGTATAAGTTGATGA 195
 Db 11320 TTAATTTAATATAATTAATTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTAAT 11379
 Qy 196 TGGCGTGGTGGTAGATTAATTTAGTCTAGGTTCAAGTTCGAATTCAGATTTGGTTAGACATTTAG 255
 Db 11380 TAATTTAATTTAATTTAATTTAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTT 11439
 Qy 256 CATTATCTTTTCTAACCATAATTAATGTAATGGAACCTTTAGGAAAAAAGAA 315
 Db 11440 AATATTTTATTTTAAATTTAATTTAATTTTAAATTTTAAATTTTAAATTTAATTTAATTTAA 11499
 Qy 316 ATCAAAATTTGAAAAATCATCCGGTGGAGTCGAGAG 352
 Db 11500 AATTAATTTAATTTAATTTAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTAAG 11536

RESULT 14

US-10-204-708-49/c
 ; Sequence 49, Application US/10204708
 ; Patent No. 6677731
 ; GENERAL INFORMATION:
 ; APPLICANT: OLEK, Alexander
 ; APPLICANT: PIEPENBROCK, Christian
 ; APPLICANT: BERLIN, Kurt
 ; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
 ; TITLE OF INVENTION: by Assessing DNA Methylation
 ; FILE REFERENCE: 5013.1012
 ; CURRENT APPLICATION NUMBER: US/10/204,708
 ; CURRENT FILING DATE: 2003-05-06
 ; PRIOR APPLICATION NUMBER: PCT/EP01/03971
 ; PRIOR FILING DATE: 2001-04-06
 ; PRIOR APPLICATION NUMBER: DE 10019058.8
 ; PRIOR FILING DATE: 2000-04-06
 ; PRIOR APPLICATION NUMBER: DE 10019173.8
 ; PRIOR FILING DATE: 2000-04-07
 ; PRIOR APPLICATION NUMBER: DE 10032529.7
 ; PRIOR FILING DATE: 2000-06-30
 ; PRIOR APPLICATION NUMBER: DE 10043826.1
 ; PRIOR FILING DATE: 2000-09-01
 ; NUMBER OF SEQ ID NOS: 98
 ; SEQ ID NO 49
 ; LENGTH: 6306
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
 US-10-204-708-49

Query Match 3.8%; Score 51.6; DB 4; Length 6306;
 Best Local Similarity 50.0%; Pred. No. 0.12;
 Matches 129; Conservative 0; Mismatches 129; Indels 0; Gaps 0;

Qy 627 TTCAATTTTCACGTGATTTTGTGGTGAACGCGTACTTGAACTTATTAATTAATAC 686
 Db 447 TTTAACTTTTAACTTAAATTTTCTTAAATACCTTATTTTATTAATTAATTAATTAAT 388
 Qy 687 ATAATTTTATAAGTTTCACCTCTTATATAATCACTATAATATATATATAGGTTTAGAATGC 746
 Db 387 ATCATTTTAAATTTAAATTTTAACTAATTTTAAATTTTAAATTTTAAATTTTAAATTTAA 328
 Qy 747 CAATTTTAAAAAAGAAATAAAAAATAAGATAAATAAGATAAATAAGATAAATAAGATAA 806
 Db 327 AAATTTTCTATCAATTTTAAAAATTTTCTTAAATTTTCTTAAATTTTAAATTTTAAATTTAT 268
 Qy 807 AAATTTGAGGGGACAAATAAATAATATGAAAGTCTATTTTAAATTTTCCATTTAGAAATTC 866
 Db 267 TTTTAAATTTCTTACATATTTTATTTATTTTCTTTTACCTTTTAAATTAATTAATTAAT 208
 Qy 867 TATTTTCCTTAAATA 884
 Db 207 TCAATACCATTTTAAATA 190

RESULT 15

US-08-446-855A-1
 ; Sequence 1, Application US/08446855A
 ; Patent No. 5849573
 ; GENERAL INFORMATION:
 ; APPLICANT: Stewart, Thomas S
 ; APPLICANT: Flores, Maria V
 ; APPLICANT: O'Sullivan, William J
 ; TITLE OF INVENTION: Nucleotide sequence encoding carbamoyl
 ; TITLE OF INVENTION: phosphate synthetase II
 ; NUMBER OF SEQUENCES: 2
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Nixon & Vanderhye PC
 ; STREET: 1100 No. 5849573th Glebe Road, 8th Floor
 ; CITY: Arlington
 ; STATE: Virginia
 ; COUNTRY: USA

```
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/446,855A
FILING DATE: 06-Jul-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Mitchard, Leonard C
REGISTRATION NUMBER: 29,009
REFERENCE/DOCKET NUMBER: 47-80
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-816-4000
TELEFAX: 703-816-4100
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 8920 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: genomic
US-08-446-855A-1

Query Match      3.8%; Score 51.6; DB 2; Length 8920;
Best Local Similarity 53.5%; Pred.No.0.13;
Matches 108; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

QY 662 GTACTTGAACTTATTATTAATTACATAATTTTATAAGTTTCACCTTCTTATATAACTC 721
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 526 GTTCCTATAATTTTATTATATATTTATTATTATTATTATTATTATTATTATTATTTC 585
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 722 ATATAATATATAGGGTTTAGAATGCCAATTTTAAAAAAGATAAAAAAATAAATAGAA 781
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 586 TTAGTTTATAAATAGTAATTTCTACTAATTTAAAAAATAAAAAAATAAAAAAATAAAA 645
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 782 TAAATCGAAAAATGAATGTAAAAATTTGAGGGCGACAAATAAATATGAAGCTCTA 841
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 646 GAAAAAATAAATTTACATATGAAAAATGAAGCTGTATATGTAATTTTATAAATATTT 705
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 842 TTATTTAAATTTTCCATTAGAA 863
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 706 AACATAAATATAAATGTATAA 727
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Search completed: February 24, 2004, 11:09:40
Job time : 138.614 secs
```

BEST AVAILABLE COPY

CC active and in a high amount suitable for medical use
XX Sequence 1428 BP; 499 A; 249 C; 208 G; 472 T; 0 U; 0 Other;
SQ

Query Match 93.8%; Score 1282.8; DB 9; Length 1428;
Best Local Similarity 95.6%; Pred. No. 2.5e-198;
Matches 1365; Conservative 0; Mismatches 2; Indels 61; Gaps 2;

QY 1 TAAATATATCTATACATTAATAAATTTGATTTTAAATTTTAGAAATTCATGATTTTATT 60
DB 1 TAAATATATCTATACATTAATAAATTTGATTTTAAATTTTAGAAATTCATGATTTTATT 60
QY 61 TTTTATACAGAAATCCGTTAATATGTTTAAATATTAACCAACTAAATTTTAAATTTTA 120
DB 61 TTTTATACAGAAATCCGTTAATATGTTTAAATATTAACCAACTAAATTTTAAATTTTA 120
QY 121 TTTTAAAGCAATTAAGCATGTTTGAATAAATATATATATGTTTAAATATCTTTTCAAAA 180
DB 121 TTTTAAAGCAATTAAGCATGTTTGAATAAATATATATGTTTAAATATCTTTTCAAAA 180
QY 181 GTATAAGTTGATGATGCGTGGTGTAGATTTATTTTAGTTCTAGGTTTCAAGTT 240
DB 181 GTATAAGTTGATGATGCGTGGTGTAGATTTATTTTAGTTCTAGGTTTCAAGTT 240
QY 241 GGTTTAGCAATTTAGCTTTATCTTTTCTTAACCAAAATTAATGTAATGGAACCTT 300
DB 241 GGTTTAGCAATTTAGCTTTATCTTTTCTTAACCAAAATTAATGTAATGGAACCTT 300
QY 301 TAGGAAAAAAGAAATCAAAATTTGAAACATCATCCGTTGAGTCGAGAGCCACACC 360
DB 301 TAGGAAAAAAGAAATCAAAATTTGAAACATCATCCGTTGAGTCGAGAGCCACACC 360
QY 361 CACGTGACCAACAATATTAATAAGAGTTTGTCTACAGTAATGCGATCTTTTATA 420
DB 361 CACGTGACCAACAATATTAATAAGAGTTTGTCTACAGTAATGCGATCTTTTATA 420
QY 421 TTCAATATCTTTTCCACTTCAAAATCTTTGAGATTTGCACCGTTTAACTTAATGTT 480
DB 421 TTCAATATCTTTTCCACTTCAAAATCTTTGAGATTTGCACCGTTTAACTTAATGTT 480
QY 481 ATATCAACGGTCTTAAATAAATCTTTGTACCGTCCCTCACATTTCAACTTTTGGCACC 540
DB 481 ATATCAACGGTCTTAAATAAATCTTTGTACCGTCCCTCACATTTCAACTTTTGGCACC 540
QY 541 TGAAAGCGGT-TATGTTTAGTTAGTTTTCGACAGTTTGAAGCGCATCTCAGGAGGC 599
DB 541 TAGAAGCGGTCTATGTTTAGTTTGTGTTGCAAGTTTGAAGCGCATCTCAGGAGGC 600
QY 600 TACTTGGTCTTGGCTTTTGGCTTTTGTTCATTTTTCACGTGATTTTGTGGTGAACAC 659
DB 601 TACTTGGTCTTGGCTTTTGGCTTTTGTTCATTTTTCACGTGATTTTGTGGTGAACAC 660
QY 660 GGGTACTTGAACCTTATTAATAATTAATATTTTATTAAGTTTCACTTCTTATATATAC 719
DB 661 GGGTACTTGAACCTTATTAATAATTAATATTTTATTAAGTTTCACTTCTTATATATAC 720
QY 720 ----- 719
DB 721 TTCATTCATGATTTATATTTTGTATGATTAATAAAGAGTTTGTAAATAATATATT 780
QY 720 TCATATATATATAGGTTTGAATGCCAATTTTAAATAAAGATTAATAATAATATAG 779
DB 781 TCATATATATATAGGTTTGAATGCCAATTTTAAATAAAGATTAATAATAATATAG 840
QY 780 AATAAAATCGAAAAATGAAATGTAATAAATTTGAGGGGGCAATTAATAATCGAAGTC 839
DB 841 AATAAAATCGAAAAATGAAATGTAATAAATTTGAGGGGGCAATTAATAATCGAAGTC 900
QY 840 TATTATTTAAATTTTCCATTTAGAAATCTTATTTTCTTGTAGTTTAAATGAGCTAGCAGTTG 899
DB 901 TATTATTTAAATTTTCCATTTAGAAATCTTATTTTCTTGTAGTTTAAATGAGCTAGCAGTTG 960
QY 900 GGAGATACACGAAATGTCATGAACAGTTGCAATGAGGAAATTAATGTAGTAGAGGA 959

DB 961 GGAGATACACGAAATGTCATGAACAGTTGTCATGTAGGAAATTAATGTAGTAGAGGA 1020
QY 960 TAGCAAGACAAAAATCCAGCCAAAGCTAGCTGCTCAGCGAACTCGATCCACAGTCTTT 1019
DB 1021 TAGCAAGACAAAAATCCAGCCAAAGCTAGCTGCTCAGCGAACTCGATCCACAGTCTTT 1080
QY 1020 TAGCAGAGTTTCAAAACGGATGAAATCTGCATGGCATGCAACTAAAGCATTTGTTCTCAGCTG 1079
DB 1081 TAGCAGAGTTTCAAAACGGATGAAATCTGCATGGCATGCAACTAAAGCATTTGTTCTCAGCTG 1140
QY 1080 CCAAGTACCCCTCACAATCAACCAACCTTTGTTTCTCCCATGATGATTTAACTCAAG 1139
DB 1141 CCAAGTACCCCTCACAATCAACCAACCTTTGTTTCTCCCATGATGATTTAACTCAAG 1200
QY 1140 TTTATCTTTCTTTGCTTTCTGAAATTTCAAGAGCTCAAAACAGCTCGAGCTCCAATCTT 1199
DB 1201 TTTATCTTTCTTTGCTTTCTGAAATTTCAAGAGCTCAAAACAGCTCGAGCTCCAATCTT 1260
QY 1200 GTGACCAACACGCGCAAAAGAAAGAAATCTCATCCGTTCACTTTAGCCACTTTAAAG 1259
DB 1261 GTGACCAACACGCGCAAAAGAAAGAAATCTCATCCGTTCACTTTAGCCACTTTAAAG 1320
QY 1260 CTAGCCAAACGGTGTATCTTCTATATATTTAGTCTCTTAACACAAACACTACCA 1319
DB 1321 CTAGCCAAACGGTGTATCTTCTATATATTTAGTCTCTTAACACAAACACTACCA 1380
QY 1320 TTATTCAATTTCAAAACCTTCTCTATATCTATACACACTAGAGAATA 1367
DB 1381 TTATTCAATTTCAAAACCTTCTCTATATCTATACACACTAGAGAATA 1428

RESULT 3
ABZ10246
ID ABZ10246 standard; DNA; 8056 BP.
XX AC ABZ10246;
XX DT 16-JAN-2003 (first entry)
XX DE Haematopoietic cell proliferation disorder related DNA sequence #386.
XX KW Human; haematopoietic cell proliferation disorder; cytostatic;
XX KW gene therapy; lymphocytic leukaemia; acute myelogenous leukaemia;
XX KW cytosine methylation state; gene; ds.
XX OS Homo sapiens.
XX PN WO200277272-A2.
XX PD 03-OCT-2002.
XX PF 26-MAR-2002; 2002WO-EP003401.
XX PR 26-MAR-2001; 2001US-0278333P.
XX PA (EPIG-) EPIGENOMICS AG.
XX PI Berlin K, Braun A, Distler J, Guetig D, Howe A, Mueller J;
XX PI Olek A, Piepenbrock C, Adorjan P, Grabs G, Lesche R, Leu E;
XX PI Lewin A, Lipscher E, Maier S, Model F, Mueller V, Otto T, Pelet C;
XX PI Schwobe I, Ziebarth H;
XX DR WPI; 2003-018942/01.
XX PT Detecting and differentiating between hematopoietic cell proliferative
XX PT disorders, comprises contacting a target nucleic acid with a reagent that
XX PT distinguishes between methylated and non-methylated CpG dinucleotides.
XX PS Claim 28; SEQ ID NO 386; 117pp; English.
XX CC The present invention describes a method for detecting and
XX CC differentiating between haematopoietic cell proliferative disorders

CC associated with at least 1 gene and/or their regulatory regions in a
 CC subject. The method comprises contacting a target nucleic acid in a
 CC biological sample obtained from the subject with at least 1 reagent,
 CC which distinguishes between methylated and non-methylated CpG
 CC dinucleotides within the target nucleic acid. ABZ09861 to ABZ11118
 CC represent specifically claimed nucleotide sequences from the present
 CC invention. Oligonucleotides from the present invention can be used: for
 CC differentiating between healthy haematopoietic cells and proliferative
 CC disorder haematopoietic cells; for differentiating between acute
 CC lymphocytic leukaemia and acute myelogenous leukaemia; as probes for
 CC determining the cytosine methylation state and/or single nucleotide
 CC polymorphisms (SNPs) of haematopoietic cell proliferation disorder
 CC related sequences and their complements; and as primers for the
 CC amplification of haematopoietic cell proliferation disorder related DNA
 CC sequences. The nucleotide sequences from the present invention can also
 CC be used for detecting a predisposition to, differentiation between
 CC subclasses, diagnosis, prognosis, treatment and/or monitoring of
 CC haematopoietic cell proliferation disorders. The present method enables a
 CC highly specific classification of haematopoietic cell proliferative
 CC disorders allowing for improved and informed treatment of patients
 XX
 SQ Sequence 8056 BP; 3711 A; 0 C; 371 G; 3974 T; 0 U; 0 Other;

Query Match 5.5%; Score 74.6; DB 7; Length 8056;
 Best Local Similarity 44.6%; Pred. No. 0.0029;
 Matches 393; Conservative 0; Mismatches 474; Indels 14; Gaps 2;

QY 9 TCTATACATTAATAAATTCATTTTAAATTTAGAAATTCATGATTTATTTTTTTTAA 68
 DB 1833 TTTTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1892
 QY 69 CCAGAAATCCGTTAATATGTGTAATAATATACCACTAATTAATAATTTTAAAGG 128
 DB 1893 TTTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1951
 QY 129 CAATTAGCATGTTGTGATAAATAATATATATGTTATATAAT-----ACTTTT 175
 DB 1952 AATATATTTTAAATTAATAATATATATATATATTTTAAATTTAAATAAATTT 2011
 QY 176 CAAAAGTATAAAGTTGATGCGGTGGGTAGATTATTTTATGTTCTAGGTCGAATGC 235
 DB 2012 TAAATATATTTAAATAAATAATATTTTAAATAATTTTATTTTAAAAAATTTTATTA 2071
 QY 236 AAGTTGGTTAGACATTTAGCCATTCTTTTTCACCAATAAATAAATAAATAAATAA 295
 DB 2072 AAAATATTTTAAATATTAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2131
 QY 296 ACCTTTAGGAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 355
 DB 2132 TTTATTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2191
 QY 356 ACACCCAGTGACCAATATTAATAAATAAAGATTGCTCTACAGTAATGCGATCATT 415
 DB 2192 AAAAAATGAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2251
 QY 416 TTTTATTTCAATACCTTTTCCACTTCTTAAATCTTGGAGATTGACCGTTAACTAATA 475
 DB 2252 TAAATATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2311
 QY 476 GTGTTATATCAACGGCTCAAAAAAATACTGTGTACGCGCTCACATTTCAACTTGGC 535
 DB 2312 TTAATAATTTTGTATTAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2371
 QY 536 CACCTGAAACCGCTATGTTTGGTTAGTGTGTTGCAACAGTTGACGGCATCACTCAGG 595
 DB 2372 ATAAATAATTTTAAATAATTTTATTTTATTTTGGAAAAATAAATAAATAAATAA 2431
 QY 596 AGGCTACTGTGCTGCTTTTTCGCTCTTTTGTTCATTTTTCACGTTATTTGTGGTGA 655
 DB 2432 TTTTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2491
 QY 656 ACACGCGTACTGAACTATATTAATAAATAAATAAATAAATAAATAAATAAATAA 715

DB 2492 TAATTAATAATTAATGAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2551
 QY 716 ATACTCATATATATATAGGCTTTAGAAATGCAATTTTAAAAAAGAATAAATAAATA 775
 DB 2552 ATTTAAAAATTTAAATTAATAATTTTAAAAAATAAATAAATAAATAAATAAATA 2611
 QY 776 ATAGAATAAATAATCGAAAAAATAAATAATGTAAAAAATTTGAGGGGACAAATAAATA 835
 DB 2612 ATATAATAATTTTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2671
 QY 836 AGTCTATTATTAAATTTTCCATTAGAAATCTATTTCCTT 876
 DB 2672 AATAAATTTTAAATTTTAAAAATTTTAAATTTTAAATTTTAAATTTTAAATTTT 2712
 RESULT 4
 ABZ10246/c
 ID ABZ10246 standard; DNA; 8056 BP.
 AC ABZ10246;
 XX
 DT 16-JAN-2003 (first entry)
 XX
 DE Haematopoietic cell proliferation disorder related DNA sequence #386.
 XX
 KW Human; haematopoietic cell proliferation disorder; cytostatic;
 KW gene therapy; lymphocytic leukaemia; acute myelogenous leukaemia;
 KW cytosine methylation state; gene; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200277272-A2.
 XX
 PD 03-OCT-2002.
 XX
 PF 26-MAR-2002; 2002WO-EP003401.
 XX
 PR 26-MAR-2001; 2001US-0278333P.
 XX
 PA (EPIG-) EPIGENOMICS AG.
 XX
 PI Berlin K, Braun A, Distler J, Guetig D, Howe A, Mueller J;
 PI Olek A, Pilepbrock C, Adorjan P, Grabs G, Lesche R, Leu B;
 PI Lewin A, Lipscher E, Maier S, Model F, Mueller V, Otto T, Pelet C;
 PI Schwöpe I, Ziebarth H;
 XX
 WPI; 2003-018942/01.

XX Detecting and differentiating between hematopoietic cell proliferative
 PT disorders, comprises contacting a target nucleic acid with a reagent that
 PT distinguishes between methylated and non-methylated CpG dinucleotides.

PS Claim 28; SEQ ID NO 386; 117pp; English.

XX The present invention describes a method for detecting and
 CC differentiating between haematopoietic cell proliferative disorders
 CC associated with at least 1 gene and/or their regulatory regions in a
 CC subject. The method comprises contacting a target nucleic acid in a
 CC biological sample obtained from the subject with at least 1 reagent,
 CC which distinguishes between methylated and non-methylated CpG
 CC dinucleotides within the target nucleic acid. ABZ09861 to ABZ11118
 CC represent specifically claimed nucleotide sequences from the present
 CC invention. Oligonucleotides from the present invention can be used: for
 CC differentiating between healthy haematopoietic cells and proliferative
 CC disorder haematopoietic cells; for differentiating between acute
 CC lymphocytic leukaemia and acute myelogenous leukaemia; as probes for
 CC determining the cytosine methylation state and/or single nucleotide
 CC polymorphisms (SNPs) of haematopoietic cell proliferation disorder
 CC related sequences and their complements; and as primers for the
 CC amplification of haematopoietic cell proliferation disorder related DNA
 CC sequences. The nucleotide sequences from the present invention can also
 CC be used for detecting a predisposition to, differentiation between
 CC subclasses, diagnosis, prognosis, treatment and/or monitoring of

[illegible]

```

QY 69  CCGAAATCGTTAATATGTTTAAATATTTACCACTAAATTTATAATTTTAAAG 128
Db 1893 TTTATAAAATGAAAAATTTTATAAAAAAATAAATTAATTTAT-AATATAAATAAAAA 1951
QY 129 CAATTAGCATGTTGTATAAATATATATATTTGTTATAAAT-----ACTTTT 175
Db 1952 AATATAATTTTATAAATAAATAATTAATAATTAATTTTATAAATTTATAAATTTT 2011
QY 176 CAAAGATATAAAGTTGATGCGGTGCGGTAGATTAATTTTAAATTTTAAATGTC 235
Db 2012 TAAATATATTTAAATAAATAATATATTTTAAATATTTTATTTTAAAAAAATTTATTA 2071
QY 236 AAGTTGGTTAGACATTTAGCCTTAATCTTTTCTTACCAAAATAAATGAATGAA 295
Db 2072 AAAATATTTTAAATATTAATAAATAAATAAATAAATAAATAAATAAATAAATTT 2131
QY 296 ACCTTTAGGAAAAAAGAAATCAAAATTCGAAACATCATCCGTCGAGTCGAGAGCC 355
Db 2132 TTTATTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2191
QY 356 ACACCCAGTCACCAATATATAAATAAAGAGTTGCTCTACGTAATGCGATCTT 415
Db 2192 AAAAAACGAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATTT 2251
QY 416 TTTTATTCATCTTTTCCACTCTTAATCTTGGAGATTTGCACGGTTAACTAATAA 475
Db 2252 TAATTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAAT 2311
QY 476 GTGTTATATCCACGGCTCAAAAAAATCTGTGTACCGTCCTCACATTTCAACTTGG 535
Db 2312 TTTATTAATTTTCGATTTAAATAAATTTTATTAATTTTAAATAAATAAATAAATTT 2371
QY 536 CACCTGAAGCCGTTAGTTAGTTAGTTGTTGCAACAGTTGAACGCGATCACTCAG 595
Db 2372 ATAAATAATTTATTAATAATTTTATTTATTTTCGAAAAATAAATAAATAAATAAAT 2431
QY 596 AGGCTACTGCTGTTGTTGCTCTTTTGTTCATTTTTCACGTTGTTTGTGCTGA 655
Db 2432 TTTTAAAAATTAATTTTATTTTATTTTATTTATTTATTTATTTATTTTATTTTAA 2491
QY 656 ACACGGTACTGAACTATATATAAATACATTAATTTTAAAGTTTCACTCTTATATA 715
Db 2492 TAAATTAATAATTAACGAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2551
QY 716 ATACTATATAATATAGGTTTATAGGTTTATAGGTTTATAGGTTTATAGGTTTATAG 775
Db 2552 ATTTAAAAATTTAAATAAATTTTAAATAAATAAATAAATAAATAAATAAATAAATAA 2611
QY 776 ATAGATAAATAATCGAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAAT 835
Db 2612 ATATAAATAATTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAAT 2671
QY 836 AGTCTATATTAATTTTCAATAGAAATTTCTATTTTCCTT 876
Db 2672 AATAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAAT 2712

```

RESULT 6

AAS46430

ID AAS46430 standard; DNA; 6106 BP.

XX AC AAS46430;

XX 18-DEC-2001 (first entry)

XX Tumour suppressor gene derived chemically modified sequence #152.

XX Human; tumour suppressor gene; oncogene; antitumour; cytostatic; cancer;

XX tumour; CpG dinucleotide; single-nucleotide polymorphism; SNP;

XX cytosine methylation; ds.

XX Homo sapiens.

OS

```

XX PN WO200168912-A2.
XX PD 20-SEP-2001.
XX PF 15-MAR-2001; 2001WO-EP002955.
XX PR 15-MAR-2000; 2000DE-01013847.
XX PR 06-APR-2000; 2000DE-01019058.
XX PR 07-APR-2000; 2000DE-01019173.
XX PR 30-JUN-2000; 2000DE-01032529.
XX PR 01-SEP-2000; 2000DE-01043826.
XX (EPIG-) EPIGENOMICS AG.
XX Olek A, Piepenbrock C, Berlin K;
XX WPI; 2001-602752/68.
XX Fragments of chemically modified genes associated with tumor suppressor
XX genes and oncogenes, useful in designing primers and probes for analyzing
XX diseases associated with cytosine methylation state e.g. cancer.
XX Claim 1; SEQ ID NO 152; 27pp; English.
XX The invention relates to a nucleic acid comprising a sequence of 18
XX bases, of a segment of chemically pretreated DNA (CP DNA) e.g. with
XX bisulphite, of genes associated with tumour suppression and oncogenes
XX having a sequence taken from 536 (actually 533 since numbers 408, 458 and
XX 500 are missing from the sequence listing) sequences (SS) and sequences
XX complementary to (SS). The nucleic acid may be a peptide nucleic acid-
XX oligomer (PNA) of at least 9 nucleotides and may form part of a set of
XX probes for detecting the cytosine methylation state and/or single
XX nucleotide polymorphisms and also to be used in an array for analysing
XX diseases associated with CpG dinucleotides e.g. cancers and tumours. The
XX probes can also be used in a method for ascertaining genetic and/or
XX epigenetic parameters for the diagnosis and/or therapy of existing
XX diseases or the predisposition to specific diseases, by analysing
XX cytosine methylations. The parameters may be compared to another set of
XX genetic and/or epigenetic parameters, the differences serving as basis
XX for diagnosis and/or prognosis events which are disadvantageous to
XX patients. The present sequence is one of the 533 genomic sequences
XX derived from tumour suppressor genes and oncogenes. Sequences with even
XX numbered Seq ID numbers are the complementary sequence of the
XX corresponding odd numbered sequence (e.g. ID 2 and ID1, ID 536 and ID
XX 535, except for those whose partner sequence is missing). Note: The
XX sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences

```

XX Sequence 6106 BP; 2270 A; 30 C; 1049 G; 2757 T; 0 U; 0 Other;

Query Match 4.7%; Score 64.4; DB 4; Length 6106;

Best Local Similarity 56.5%; Pred. No. 0.13;

Matches 139; Conservative 0; Mismatches 106; Indels 1; Gaps 1;

```

QY 1 TAAATAATCTATACATTAATAAATTTGATTTTAAATTTTAGAAATTCATGATTTTAT 60
Db 797 TATTTTAAAGTTTATTTTATTTAGAGTTAATTAATAAATTCGCTAAATAATTTTATTTT 856
QY 61 TTTTATTTACCAGAAATCCGTTAATAATCTGTTAAATATTTACCACTAAATTTATAATTTTA 120
Db 857 TTTTATTTTATTTATTAATTTTATTTTATTTAGATGAATTAAGAAATTTTATTTTAA 916
QY 121 TTTTAAAGCAATTAAGCATGTTT-GATAAAATATATATATTTGTTATAAATACTTTTCAA 179
Db 917 ATTAATTTGATTTGTAATTTTATTTTATTTAGATGAATTAAGAAATTTAGATTTGTA 976
QY 180 AGTATAAGTTGATGATGCGGTGGTGGTAGATTTATTTAGTTCTAGTTTCAATGCAAGT 239
Db 977 AATATAAATAAATAGAAAGATTTGTTATTTATTTATTTTATTTTATTTTGGTTTGAAGT 1036
QY 240 TGGTTT 245

```

```
Db 1037 GAGTAT 1042
|||
797 TATTTAAAGTTTATTTAGAGTTAAATAAATAATGCGTAAAAAATTTATTTT 856
QY 61 TTTTTCACGAAATCCGTTAATATTGTTAAATATTACCACTAATTATAAATTTA 120
|||
857 TTTTTCATTTATTTAATTTTATTTTATTCAGATGAATTAAGAAATTTTAAATTTA 916
QY 121 TTTTAAGCAATTAAGCATGTTT-GATAAAATATATATTTTATATAACTTTTCAA 179
Db 917 ATTAATTTGATTTGTAATTTTATTTAGTTAAGATAATATATTATAGATTAGATTGTAA 976
QY 180 AGTATAAGTTGATGATGGCGTGGTGTAGATTATTTTATTTAGTTTCGAATGCAAGT 239
Db 977 AATATAAATAATTAGAAAGAGTATTTGTTATTTTATTTTATTTTATTTTGGTTGTTTGAAGT 1036
QY 240 TGGTTT 245
Db 1037 GAGTAT 1042

RESULT 8
ABL33473
ID ABL33473 standard; DNA; 6106 BP.
XX AC ABL33473;
XX AC
XX DT 26-MAR-2002 (first entry)
XX DE Human immune system associated gene SEQ ID NO: 1446.
XX KW Human; immune system disease; cytosine methylation; antiasthmatic;
KW antiarteriosclerotic; antianaemic; cytostatic; nootropic;
KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
KW antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
KW antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;
XX ds.
XX OS Homo sapiens.
XX PN WO200200928-A2.
XX PD 03-JAN-2002.
XX PF 02-JUL-2001; 2001WO-EP007537.
XX PR 30-JUN-2000; 2000DE-01032529.
XX PR 01-SEP-2000; 2000DE-01043826.
XX PA (EPIG-) EPIGENOMICS AG.
XX PI Olek A, Piepenbrock C, Berlin K;
XX WPI; 2002-130909/17.
XX Nucleic acid comprising fragment of chemically modified gene, useful for
XX diagnosis and treatment of diseases associated with abnormal cytosine
XX methylation.
XX Claim 1; SEQ ID NO 1446; 32pp + Sequence Listing; German.
XX The present invention provides a number of human immune system associated
XX genes which are modified by the methylation of cytosines. The sequences
XX can be used in the diagnosis and treatment of immune system disorders,
XX including eye diseases such as retinopathy, neovascular glaucoma and
XX macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
XX leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
XX rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
XX diseases. The present sequence is a gene of the invention
XX
XX Sequence 6106 BP; 2270 A; 30 C; 1049 G; 2757 T; 0 U; 0 Other;
```

```
Db 1037 GAGTAT 1042
|||
797 TATTTAAAGTTTATTTAGAGTTAAATAAATAATGCGTAAAAAATTTATTTT 856
QY 61 TTTTTCACGAAATCCGTTAATATTGTTAAATATTACCACTAATTATAAATTTA 120
|||
857 TTTTTCATTTATTTAATTTTATTTTATTCAGATGAATTAAGAAATTTTAAATTTA 916
QY 121 TTTTAAGCAATTAAGCATGTTT-GATAAAATATATATTTTATATAACTTTTCAA 179
Db 917 ATTAATTTGATTTGTAATTTTATTTAGTTAAGATAATATATTATAGATTAGATTGTAA 976
QY 180 AGTATAAGTTGATGATGGCGTGGTGTAGATTATTTTATTTAGTTTCGAATGCAAGT 239
Db 977 AATATAAATAATTAGAAAGAGTATTTGTTATTTTATTTTATTTTATTTTGGTTGTTTGAAGT 1036
QY 240 TGGTTT 245
Db 1037 GAGTAT 1042

RESULT 7
ABK40032
ID ABK40032 standard; DNA; 6106 BP.
XX AC ABK40032;
XX AC
XX DT 21-MAY-2002 (first entry)
XX DE Human chemically pretreated gene sequence #57 strand 2.
XX KW Human; ds; bisulphite treatment; CpG; DNA methylation; cancer; tumour;
KW cytostatic; ALDH6; CYP11A; CYP3A3; DYPD; EPHX2; OCLN; TXNRD1;
KW UGT8; MRP; pharmacogenomics; SNP; single nucleotide polymorphism.
XX OS Homo sapiens.
XX PN WO200202806-A2.
XX PD 10-JAN-2002.
XX PF 29-JUN-2001; 2001WO-BP007470.
XX PR 30-JUN-2000; 2000DE-01032529.
XX PR 01-SEP-2000; 2000DE-01043826.
XX PA (EPIG-) EPIGENOMICS AG.
XX PI Olek A, Piepenbrock C, Berlin K;
XX WPI; 2002-154757/20.
XX New nucleic acid, oligonucleotides and peptide nucleic acid-oligomers,
XX useful for detecting cytosine methylation state of genes associated with
XX pharmacogenomics and for therapy of diseases e.g. cancer.
XX
XX Claim 1; SEQ ID NO 114; 24pp; English.
XX The invention relates to a nucleic acid comprising a sequence at least 18
XX bases in length of a segment of the chemically pretreated DNA of genes
XX associated with pharmacogenomics according to one of the sequences of the
XX genes ALDH6 (NM 000693), CYP11A (NM 000781), CYP11B1 (NM 000497), CYP3A3
XX (NM 000776 and NM 017460), DYPD (NM 000110), EPHX2 (NM 001979), OCLN
XX (NM 002538), TXNRD1 (NM 003330), UGT8 (NM 003360), MRP (NM 004996),
XX NM 019900, NM 019901, NM 019902, NM 019852, NM 019898, NM 019899, and
XX their complementary sequences, or a sequence (S1) chosen from 87
XX sequences and their complements, the chemical pretreatment is bisulphite
XX treatment to convert cytosines (but not methyl-cytosines) into uracils.
XX Also included are an oligomer (II) in particular an oligonucleotide or a
XX peptide nucleic acid (PNA)-oligomer, comprising in each case at least one
XX base sequence having a length of 9 nucleotides which hybridises to or is
XX identical to a chemically pretreated DNA of genes associated with
XX pharmacogenomics and their complements, arranged in an array for
XX analysing diseases associated with the methylation state (CpG) and/or
XX detecting SNPs (single nucleotide polymorphisms) of the 87 sequences. The
XX oligomers may also be used as PCR primers. The set of 87 nucleic acids
XX and their complements is useful for diagnosis and therapy of solid
XX tumours and cancer. The present sequence represents one the 87 DNA
XX sequences or its complement. Note: The sequence data for this patent did
XX not form part of the printed specification, but was obtained in
XX electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 6106 BP; 2270 A; 30 C; 1049 G; 2757 T; 0 U; 0 Other;
```

```
Query Match 4.7%; Score 64.4; DB 6; Length 6106;
Best Local Similarity 56.5%; Pred.No.0.13;
Matches 139; Conservative 0; Mismatches 106; Indels 1; Gaps 1;
QY 1 TAAATAATCTATACATTAAAAAATTTGATTTTAAATAATTTAGAAATTCATGATTTTAT 60
```


XX The invention relates to an isolated nucleic acid molecule (I) comprising
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
 CC specification or its complement. (I) is useful for: (a) assessing whether
 CC a patient is afflicted with prostate cancer; (b) monitoring the
 CC progression of prostate cancer in a patient; (c) assessing the efficacy
 CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
 CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;
 CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
 CC determining whether prostate cancer has metastasized in a patient; (h)
 CC assessing the aggressiveness or indolence of prostate cancer in a patient
 CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
 XX
 SQ Sequence 309 BP; 188 A; 13 C; 11 G; 97 T; 0 U; 0 Other;
 Query Match 4.7%; Score 64; DB 5; Length 309;
 Best Local Similarity 56.0%; Pred. No. 0.14;
 Matches 121; Conservative 0; Mismatches 95; Indels 0; Gaps 0;
 QY 621 CTTTCTGCAATTTTTCACGTGATTTTGTGGTGAACACGGTACTTGAACCTTATATA 680
 Db 19 CTTTCTGCAATTTTTCACGTGATTTTGTGGTGAACACGGTACTTGAACCTTATATA 78
 QY 681 AATTACATATTTTATAAGTTTCACCTCTATATATACTCATATATATATAGGTTTA 740
 Db 79 TTTTCTGCAATTTTTCACGTGATTTTGTGGTGAACACGGTACTTGAACCTTATATA 138
 QY 741 GAATGCCAATTTTAAAAAAGAAATAAAAAATAAATAAGATAAATCGAAAAATGAAA 800
 Db 139 AAAAAAATTTTGAAGAAAAAATTTTGAAGAAAAAATTTTGAAGAAAAAATTTTGAAG 198
 QY 801 TGTAAAAAATTTTGAAGAAAAAATTTTGAAGAAAAAATTTTGAAGAAAAAATTTTGAAG 836
 Db 199 AAAAAAATTTTGAAGAAAAAATTTTGAAGAAAAAATTTTGAAGAAAAAATTTTGAAG 234
 RESULT 11
 AAS45325/c
 ID AAS45325 standard; DNA; 7195 BP.
 XX
 AC AAS45325;
 XX
 DT 18-DEC-2001 (first entry)
 XX
 DE Chemically pretreated complementary DNA associated with cell cycle #15.
 XX
 KW Cell cycle; human; CpG dinucleotide; cytosine methylation; HIV; aging;
 KW human immunodeficiency virus; neurodegenerative disorder; solid tumour;
 KW graft-versus-host disease; glomerular disease; Lewy body disease; cancer;
 KW arthritis; arteriosclerosis; anti-HIV; neuroprotective; antiarthritic;
 KW immunosuppressive; antitumour; cytostatic; antiarteriosclerotic; ds;
 KW PCR primer.
 XX
 OS Homo sapiens.
 XX
 PN WO200168911-A2.
 XX
 PD 20-SEP-2001.
 XX
 PF 15-MAR-2001; 2001WO-EP002945.
 XX
 PR 15-MAR-2000; 2000DE-01013847.
 PR 06-APR-2000; 2000DE-01019058.
 PR 07-APR-2000; 2000DE-01019173.
 PR 30-JUN-2000; 2000DE-01032529.
 PR 01-SEP-2000; 2000DE-01043826.
 XX
 PA (EPIC-) EPICENOMICS AG.
 XX
 PI Olek A, Piepenbrock C, Berlin K;
 XX
 DR WPI; 2001-602751/68.

XX Designing primers and probes for analyzing diseases associated with
 PT cytosine methylation state e.g. arthritis, cancer, aging,
 PT arteriosclerosis comprising fragments of chemically modified genes
 PT associated with cell cycle.
 XX
 PS Claim 1; SEQ ID NO 30; 28pp; English.
 XX
 CC Sequences AAS45296-AAS45520 represent chemically pretreated genomic DNA
 CC molecules associated with the cell cycle and specific PCR primers of the
 CC invention. The sequences are useful for detecting the methylation state
 CC of all CpG dinucleotides in a sequence and therefore for analysing
 CC associated diseases. By analysing cytosine methylations in the pretreated
 CC DNA, genetic and/or epigenetic parameters for the diagnosis and therapy
 CC of existing diseases or the predisposition to specific diseases can be
 CC ascertained. The parameters may be compared to another set of genetic
 CC and/or epigenetic parameters, the differences serving as basis for
 CC diagnosis and/or prognosis events which are disadvantageous to patients.
 CC The sequences of the invention are useful for the diagnosis and therapy
 CC of HIV infection, neurodegenerative disorders, graft-versus-host disease,
 CC aging, glomerular disease, Lewy body disease, arthritis,
 CC arteriosclerosis, solid tumours and cancers
 XX
 SQ Sequence 7195 BP; 1873 A; 224 C; 1538 G; 3560 T; 0 U; 0 Other;
 Query Match 4.6%; Score 63.4; DB 4; Length 7195;
 Best Local Similarity 50.2%; Pred. No. 0.18;
 Matches 157; Conservative 0; Mismatches 156; Indels 0; Gaps 0;
 QY 675 ATTATAAATACATAAATTTTATAAGTTTCACCTCTTATATAACTCATATATATATAG 734
 Db 6759 ATTATCCAAAAACAATCAITTTTAATTTAATTTTATTTTAACACATATATTTTAAT 6700
 QY 735 GGTTAGAATGCCAATTTTAAAAAAGAAATAAAAAATAAATAAGATAAATCGAAAA 794
 Db 6699 TATTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 6640
 QY 795 ATGAAATGTAAAAAATTTTGGGGGACAAATATAAATATGAAAGTCTATTTTAAATTTT 854
 Db 6639 AAAACATTTTAAATTTTTCAAAAAACAATAAATAAATAAATAAATAAATAAATAA 6580
 QY 855 CCATTAGAATTTCTATTTTCCCTTAGTTAATATAGTACGCCAGTTGGGAGATACGAAAA 914
 Db 6579 CCAAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 6520
 QY 915 TGTCTAGAACAGTTGCGATGTGCGAATTAATCTAGTAGGGATAGCAGACAAAAT 974
 Db 6519 ACTAACGAATCAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 6460
 QY 975 CCAAGCCCAAGCTA 987
 Db 6459 ACCCAACAATAAATA 6447
 RESULT 12
 ABK28166/c
 ID ABK28166 standard; DNA; 7195 BP.
 XX
 AC ABK28166;
 XX
 DT 23-APR-2002 (first entry)
 XX
 DE DNA transcription associated complementary genomic DNA #20.
 XX
 KW DNA transcription associated gene; peptide nucleic acid; PNA-oligomer;
 KW PNA; cytosine methylation state; SNP; retroviral infection; gene; ds;
 KW single nucleotide polymorphism; adenosine deaminase deficiency; cancer;
 KW viral infection; Sezary syndrome; haematological disorder; tuberculosis;
 KW immunological disorder; Werner syndrome; developmental disorder;
 KW psoriasis; Rieger's syndrome; neurological disorder; erythropoiesis;
 KW neurodegenerative disorder; Waardenburg syndrome; Niemann-Pick disease;
 KW myelodysplastic syndrome; myocardial infarction; hypertension; arthritis;
 KW angiogenesis; congenital heart disease; HDR syndrome; gene therapy;


```
PR 11-JAN-2000; 2000US-00480902.
XX (BYAT/) BYATT J C.
PA (MATH/) MATHIALAGAN N.
PA (TAON/) TAO N.
PA (WARR/) WARREN W C.
XX
PI Byatt JC, Mathialagan N, Tao N, Warren WC;
XX
DR WPI; 2003-110599/10.
XX
PT New nucleic acid associated with lactation, and muscle and fat
PT deposition, useful for genome mapping, gene identification and analysis,
PT cattle breeding, or for genetically improving cattle.
XX
PS Claim 2; SEQ ID NO 15014; 245pp; English.
XX
CC The invention relates to a purified nucleic acid molecule associated with
CC lactation or muscle and fat deposition (designated LMPD), derived from
CC cattle, and the LMPD nucleic acid can specifically hybridise to a second
CC nucleic acid molecule comprising any of 15112 nucleotide sequences,
CC appearing as ABX34836-ABX49947, or complements of them. Also included are
CC ; (1) a transformed cell having a nucleic acid comprising an LMPD nucleic
CC acid linked to a promoter and a 3' non-translated sequence that
CC functions in the cell to cause termination of transcription and addition
CC of polyadenylated ribonucleotides to a 3' end of the mRNA molecule; and
CC (2) determining a level or pattern of a molecule in a bovine cell or
CC tissue comprising: (a) incubating a marker nucleic acid (comprising any
CC of the 15112 nucleic acid sequences or its complement or fragment) with a
CC complementary nucleic acid molecule obtained from the bovine cell or
CC tissue, where hybridisation between the marker nucleic acid and the
CC complementary nucleic acid permits the detection of the molecule; and (b)
CC detecting the level or pattern of the complementary nucleic acid, where
CC the detection of the complementary nucleic acid is predictive of the
CC level or pattern of the molecule. The LMPD nucleic acid is used for
CC determining a level or pattern of a molecule in a bovine cell or tissue.
CC It is useful for genome mapping, gene identification and analysis, cattle
CC breeding, preparation of constructs for use in cattle gene expression, or
CC for genetically improving cattle. The present sequence is one of the
CC 15112 bovine LMPD EST (expressed sequence tag) nucleic acids. Note: The
CC present sequence was not shown in the specification but was obtained in
CC electronic format from the USPTO web site:
CC seqdata.uspto.gov/sequence.html?DocID=20020137139
XX
SQ Sequence 375 BP; 143 A; 14 C; 24 G; 194 T; 0 U; 0 Other;
```

Query Match 4.6%; Score 62.2; DB 7; Length 375;
Best Local Similarity 53.0%; Pred. No. 0.28;
Matches 133; Conservative 0; Mismatches 118; Indels 0; Gaps 0;

| | | | |
|----|-----|--|-----|
| QY | 619 | GTCCTTTGTCACATTTTTCACCTGATTTTGTGGTGAACACGCTACTTGAACCTTATTA | 678 |
| Db | 276 | GTATTTTTTAAATTTTTTTTTTTTATTTTTTTTCCATCCTTTTTTTTAAAAAAA | 217 |
| QY | 679 | TAAATACATAATTTTAAAGTTTCACTCTTATATATATCTCATATATATAGGGTT | 738 |
| Db | 216 | TATATATTTTTTTTTTTTTTTTTTTTTTTTAAATAAATAAATAAATAAATAAATA | 157 |
| QY | 739 | TAGATGCCAATTTTTTAAAAAAGAAATAAAAAATAAATAAATAAATAAATAAATAA | 798 |
| Db | 156 | AATAAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA | 97 |
| QY | 799 | AATGTAAAAAATTTGAGGGGCAATAAATAATGAAAGTCTATTATTATTTCAT | 858 |
| Db | 96 | AAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA | 37 |
| QY | 859 | TAGATTCTAT | 869 |
| Db | 36 | AAAAATTAT | 26 |

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 24, 2004, 07:14:56 ; Search time 6100.31 Seconds
(without alignments)
9712.612 Million cell updates/sec

Title: US-09-743-823-21

Perfect score: 1367

Sequence: 1 taaataatctatacattaa.....ctacacacactagaagaata 1367

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_hg.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vt.*

15: em_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_mu.*

20: em_on.*

21: em_or.*

22: em_ov.*

23: em_ph.*

24: em_pl.*

25: em_pi.*

26: em_ro.*

27: em_sts.*

28: em_un.*

29: em_vt.*

30: em_hg_hum.*

31: em_hg_inv.*

32: em_hg_other.*

33: em_hg_mus.*

34: em_hg_pln.*

35: em_hg_rtd.*

36: em_hg_mam.*

37: em_hg_vrt.*

38: em_sy.*

39: em_hgo_hum.*

40: em_hgo_mus.*

41: em_hgo_other.*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | DB | ID | Description |
|------------|--------|---------------|--------|----|-------------|---------------------|
| 1 | 1367 | 100.0 | 1367 | 6 | AX006477 | AX006477 Sequence |
| 2 | 1282.8 | 93.8 | 1428 | 6 | AX840232 | AX840232 Sequence |
| 3 | 1282.8 | 93.8 | 3393 | 8 | SOXSBG78 | DL16107 Glycine max |
| c 4 | 82.4 | 6.0 | 56152 | 2 | AC116963 | AC116963 Dictyoste |
| 5 | 82.4 | 6.0 | 183648 | 3 | AC117076 | AC117076 Dictyoste |
| c 6 | 81.6 | 6.0 | 129240 | 9 | AC084128 | AC084128 Homo sapi |
| 7 | 77.4 | 5.7 | 76568 | 3 | MBREV | AF538053 Monosiga |
| c 8 | 76.4 | 5.6 | 176898 | 9 | AC117569 | AC117569 Homo sapi |
| c 9 | 76.2 | 5.6 | 923 | 3 | AY442221 | AY442221 Drosophila |
| c 10 | 76.2 | 5.6 | 12029 | 3 | AE001427 | AE001427 Plasmodiu |
| c 11 | 75 | 5.5 | 266544 | 3 | AC116956 | AC116956 Dictyoste |
| 12 | 74.6 | 5.5 | 8056 | 6 | AX599046 | AX599046 Sequence |
| c 13 | 74.6 | 5.5 | 154071 | 3 | AC115598 | AC115598 Dictyoste |
| 14 | 73.6 | 5.4 | 110000 | 3 | AC116305_3 | Continuation (4 of |
| 15 | 73 | 5.3 | 101034 | 9 | AC096550 | AC096550 Homo sapi |
| c 16 | 72.6 | 5.3 | 8056 | 6 | AX599046 | AX599046 Sequence |
| 17 | 72.6 | 5.3 | 251448 | 3 | AE014819 | AE014819 Plasmodiu |
| c 18 | 72.2 | 5.3 | 27694 | 8 | HASMT | D31785 Pichia cana |
| 19 | 72 | 5.3 | 254733 | 3 | AC117075 | AC117075 Dictyoste |
| 20 | 71.4 | 5.2 | 8056 | 6 | AX598900 | AX598900 Sequence |
| c 21 | 71.2 | 5.2 | 181792 | 9 | AC098822 | AC098822 Homo sapi |
| c 22 | 71 | 5.2 | 191542 | 2 | BX649407 | BX649407 Danio rer |
| 23 | 70.6 | 5.2 | 110000 | 2 | PFMAL7P1_00 | AL844506 Plasmodiu |
| 24 | 70.2 | 5.1 | 86826 | 3 | PFMAL3P5 | AL034556 Plasmodiu |
| c 25 | 70.2 | 5.1 | 302156 | 3 | AC116977 | AC116977 Dictyoste |
| c 26 | 69.6 | 5.1 | 72243 | 9 | AL731858 | AL731858 Human DNA |
| 27 | 69.4 | 5.1 | 252632 | 3 | AE014818 | AE014818 Plasmodiu |
| c 28 | 69 | 5.0 | 250195 | 3 | AE014831 | AE014831 Plasmodiu |
| c 29 | 68.8 | 5.0 | 163443 | 2 | AC006280 | AC006280 Plasmodiu |
| 30 | 68.8 | 5.0 | 205429 | 2 | AC005506 | AC005506 Plasmodiu |
| c 31 | 68.8 | 5.0 | 253132 | 3 | AE014846 | AE014846 Plasmodiu |
| c 32 | 68.4 | 5.0 | 7011 | 8 | DCAJ3139 | AE003139 Dausus ca |
| 33 | 68.2 | 5.0 | 57538 | 3 | AC115682 | AC115682 Dictyoste |
| c 34 | 68.2 | 5.0 | 110000 | 2 | PFMAL6P1_10 | Continuation (11 o |
| c 35 | 68.2 | 5.0 | 110000 | 2 | PFMAL7P1_07 | Continuation (8 of |
| 36 | 68.2 | 5.0 | 252394 | 3 | AE014833 | AE014833 Plasmodiu |
| 37 | 68 | 5.0 | 154995 | 9 | AC011979 | AC011979 Homo sapi |
| c 38 | 67.8 | 5.0 | 6591 | 8 | YSCMTCG06 | L36890 Saccharomyc |
| c 39 | 67.8 | 5.0 | 110000 | 2 | PFMAL7P1_06 | Continuation (7 of |
| 40 | 67.8 | 5.0 | 348600 | 1 | AB063521 | AB063521 Wiggleswo |
| 41 | 67.4 | 4.9 | 12029 | 3 | AE001373 | AE001373 Plasmodiu |
| c 42 | 67.4 | 4.9 | 34996 | 3 | U87145 | U87145 Toxoplasma |
| c 43 | 67.4 | 4.9 | 132254 | 3 | AC116330 | AC116330 Dictyoste |
| c 44 | 67.4 | 4.9 | 132254 | 3 | AC116330 | AC116330 Dictyoste |
| 45 | 67.4 | 4.9 | 253001 | 3 | AE014834 | AE014834 Plasmodiu |

ALIGNMENTS

RESULT 1
AX006477
LOCUS
DEFINITION Sequence 22 from Patent WO0004146.
ACCESSION AX006477
VERSION AX006477.1 GI:9994604
KEYWORDS
SOURCE Glycine max (soybean)
ORGANISM Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
REFERENCE 1
AUTHORS Fogher, C.

Pred. No. is the number of results predicted by chance to have a

TITLE A synthetic polynucleotide coding for human lactoferrin, vectors,
cells and transgenic plants containing it
JOURNAL Patent: WO 0004146-A 22 JAN-2000;
FOGHER CORRADO (IT); PLANTECHNO SRL (IT)

FEATURES
source Location/Qualifiers
1..1367
/organism="Glycine max"
/mol_type="unassigned DNA"
/db_xref="taxon:3847"
promoter 1..1367
/note="sequence of the promoter of the gene coding for 7S
basic globulin"

ORIGIN
Query Match 100.0%; Score 1367; DB 6; Length 1367;
Best Local Similarity 100.0%; Pred. No. 2.6e-202;
Matches 1367; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TAAATTAATCTATACATTAATAATTTGATTTTAAATTTAGAAATTCATGATTTTATT 60
Db 1 TAAATTAATCTATACATTAATAATTTGATTTTAAATTTTAAATTTTAAATTTTAAATTTTATT 60
Qy 61 TTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 120
Db 61 TTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 120
Qy 121 TTTTAAAGGCAATTAAGCATGTTTGATAAATATATATATTTGTTATTAATTAATTAATTAATTA 180
Db 121 TTTTAAAGGCAATTAAGCATGTTTGATAAATATATATATTTGTTATTAATTAATTAATTAATTA 180
Qy 181 GTATAAAGTTGATGATCGGTGGTGGTAGATTTATTTAGTTCTGATGTTGCAATGCAAGTT 240
Db 181 GTATAAAGTTGATGATCGGTGGTGGTAGATTTATTTAGTTCTGATGTTGCAATGCAAGTT 240
Qy 241 GGTTTAGACATTTAGCCTTATCTTTTCTTAACCAAAATTAATGTAATGGAACACCTT 300
Db 241 GGTTTAGACATTTAGCCTTATCTTTTCTTAACCAAAATTAATGTAATGGAACACCTT 300
Qy 301 TAGAAAAAAGAAAGAAATCRAAATTAAGAAATCATATCGGTGGAGTCGGAAGCCACACC 360
Db 301 TAGAAAAAAGAAAGAAATCRAAATTAAGAAATCATATCGGTGGAGTCGGAAGCCACACC 360
Qy 361 CACGTGACCCAAACAATTAATAATAGAGTTTGCTCTACAGTAATCGCATACTTTTTTA 420
Db 361 CACGTGACCCAAACAATTAATAATAGAGTTTGCTCTACAGTAATCGCATACTTTTTTA 420
Qy 421 TTCAATTAATTTTCCACTTTAAATTTGAGATTTGAGATTTGACCCGTTAACTAAATTAAGTT 480
Db 421 TTCAATTAATTTTCCACTTTAAATTTGAGATTTGAGATTTGACCCGTTAACTAAATTAAGTT 480
Qy 481 ATATCCAAACGGTCTTAAATAATCTGTAACCGTCTCAATTTCAACTTTTCGCAACC 540
Db 481 ATATCCAAACGGTCTTAAATAATCTGTAACCGTCTCAATTTCAACTTTTCGCAACC 540
Qy 541 TGAAGCCGTTATGTTTGTAGTTAGTTGTTGCAACAGTTGAGCGCATCTACAGAGGCT 600
Db 541 TGAAGCCGTTATGTTTGTAGTTAGTTGTTGCAACAGTTGAGCGCATCTACAGAGGCT 600
Qy 601 ACTTGGCTTCTGCTTTTCGCTTTTGTTCATTTTTCACGTGATTTTGTGTTGGAACAG 660
Db 601 ACTTGGCTTCTGCTTTTTCGCTTTTGTTCATTTTTCACGTGATTTTGTGTTGGAACAG 660
Qy 661 CGTACTTGAACACTTATTAATAATTAATAATTTTATAAGTTTCACCTCTTATATAATACT 720
Db 661 CGTACTTGAACACTTATTAATAATTAATAATTTTATAAGTTTCACCTCTTATATAATACT 720
Qy 721 CATATATATATAGGTTTGAATGCAATTTTAAATAAGTAATAATAATAATAATAATAATA 780
Db 721 CATATATATATAGGTTTGAATGCAATTTTAAATAAGTAATAATAATAATAATAATAATA 780
Qy 781 ATAAATCGAAAAAATGAATGTAATAAATTTTGAGGGGGAACAATAAATATGAAAGTCT 840
Db 781 ATAAATCGAAAAAATGAATGTAATAAATTTTGAGGGGGAACAATAAATATGAAAGTCT 840

Qy 841 ATTATTTAAATTTTCCATTTAGAAATTTCTATTTCTTCTAGTTAATATGAGTAGCCAGTTGG 900
Db 841 ATTATTTAAATTTTCCATTTAGAAATTTCTATTTCTTCTAGTTAATATGAGTAGCCAGTTGG 900
Qy 901 GAGATACAGAAAAATGTCATGAAAACAGTTGATGATGAGGAAATTAATGTTAGTAGAGGAT 960
Db 901 GAGATACAGAAAAATGTCATGAAAACAGTTGATGATGAGGAAATTAATGTTAGTAGAGGAT 960
Qy 961 AGCAAGCAAAAAATTCCAAGCCAAAGCTAGCTGCTCACCGAACTCGATCCACACGCTCTTT 1020
Db 961 AGCAAGCAAAAAATTCCAAGCCAAAGCTAGCTGCTCACCGAACTCGATCCACACGCTCTTT 1020
Qy 1021 ACAGAGTTTCAAAACGGATGAAATCTGATGCGATGCAACTTAAGCATTTCTCAGCTGC 1080
Db 1021 ACAGAGTTTCAAAACGGATGAAATCTGATGCGATGCAACTTAAGCATTTCTCAGCTGC 1080
Qy 1081 CAAGTACCCCTTCACACTCACCACCCCTTTGTTTCTCCCATTTGATGTTAACTCAAGT 1140
Db 1081 CAAGTACCCCTTCACACTCACCACCCCTTTGTTTCTCCCATTTGATGTTAACTCAAGT 1140
Qy 1141 TTATCCTTTCTTTGCTTTCTGAAAATTTCAAGGCTCAAAACAGCTCGACGTCCTCAATTTG 1200
Db 1141 TTATCCTTTCTTTGCTTTCTGAAAATTTCAAGGCTCAAAACAGCTCGACGTCCTCAATTTG 1200
Qy 1201 TGACCAACACGGCCAAAGAAAGAGATCTCTCCGTTTACACTTTAGCCACTTAAAGC 1260
Db 1201 TGACCAACACGGCCAAAGAAAGAGATCTCTCCGTTTACACTTTAGCCACTTAAAGC 1260
Qy 1261 TAGCCAAACGGTGATCTTTCTCTATATATTGATGTTAGTCTCTTAACCAACCAACACTACCAT 1320
Db 1261 TAGCCAAACGGTGATCTTTCTCTATATATTGATGTTAGTCTCTTAACCAACCAACACTACCAT 1320
Qy 1321 TATTCAATTTCAAAACCTTGCTCTATACTACACACTAGAGAATA 1367
Db 1321 TATTCAATTTCAAAACCTTGCTCTATACTACACACTAGAGAATA 1367

RESULT 2
AX840232 1428 bp DNA linear PAT 16-DEC-2003
LOCUS Sequence 6 from Patent WO03073839.
DEFINITION AX840232
ACCESSION AX840232
VERSION AX840232.1 GI:39978632
KEYWORDS Glycine max (soybean)
SOURCE Glycine max
ORGANISM Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
REFERENCE 1
AUTHORS Fogher, C. and Reggi, S.
TITLE Expression of lysosomal enzymes in plant seeds
JOURNAL Patent: WO 03073839-A 6 12-SEP-2003;
FEATURES Location/Qualifiers
1..1428
/organism="Glycine max"
/mol_type="unassigned DNA"
/db_xref="taxon:3847"
promoter 1..1428
ORIGIN
Query Match 93.8%; Score 1282.8; DB 6; Length 1428;
Best Local Similarity 95.6%; Pred. No. 2.8e-189;
Matches 1365; Conservative 0; Mismatches 2; Indels 61; Gaps 2;
Qy 1 TAAATATCTATACATTAATAATTTGATTTTAAATTTTAAATTTTAAATTTTAAATTTTATT 60
Db 1 TAAATATCTATACATTAATAATTTGATTTTAAATTTTAAATTTTAAATTTTAAATTTTATT 60
Qy 61 TTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 120

Db 61 TTTTACCAGAAATCCGTTAAATATGTTAAATATTACCACTAAATTTAATTTA 120
Qy 121 TTTTAAGCAATTAAGCATGTTGTAATAAATATATATATGTTATAAATCTTTTCAAAA 180
Db 121 TTTTAAAGCAATTAAGCATGTTGTAATAAATATATATATGTTATAAATCTTTTCAAAA 180
Qy 181 GTATAAAGTTGATGATGCGTGGTGGTGAATATATTTTAGTTCTAGTTTCGAATGCAAGTT 240
Db 181 GTATAAAGTTGATGATGCGTGGTGGTGAATATATTTTAGTTCTAGTTTCGAATGCAAGTT 240
Qy 241 GGTTTAGACATTTAGCCTTATCTTTTCTTAAACCAAAATAAATGTAATGAAACCTT 300
Db 241 GGTTTAGACATTTAGCCTTATCTTTTCTTAAACCAAAATAAATGTAATGAAACCTT 300
Qy 301 TAGGAAAAAAGAAATCAAAATTCGAAACATCATCCGTTGGAGTCGAGAGCCACACC 360
Db 301 TAGGAAAAAAGAAATCAAAATTCGAAACATCATCCGTTGGAGTCGAGAGCCACACC 360
Qy 361 CACGTGACCCAAATATTTAAATTAAGAGTTTGCTCTACAGTAATGCGATACTTTTTA 420
Db 361 CACGTGACCCAAATATTTAAATTAAGAGTTTGCTCTACAGTAATGCGATACTTTTTA 420
Qy 421 TTCATATCTTTTCCACTCTTAATCTTGAAGATTTCGACCGTTAACTAATTAAGTTT 480
Db 421 TTCATATCTTTTCCACTCTTAATCTTGAAGATTTCGACCGTTAACTAATTAAGTTT 480
Qy 481 ATATCCAAACGGTCTTAAAAAATCTTGCTACCGTCCCTCACATTTCAACTTTGCGCACC 540
Db 481 ATATCCAAACGGTCTTAAAAAATCTTGCTACCGTCCCTCACATTTCAACTTTGCGCACC 540
Qy 541 TGAAGCCGT-TATGTTAGGTAGTGTGTTGCAACAGTTGAAGCGCATCATCTAGGAGC 599
Db 541 TGAAGCCGTCTATGTTAGGTAGTGTGTTGCAACAGTTGAAGCGCATCATCTAGGAGC 600
Qy 600 TACTTGGTCTTCTGCTTTTGGCTTTTGTTCATTTTTCACGTTGTTGTTGTTGAACAC 659
Db 601 TACTTGGTCTTCTGCTTTTGGCTTTTGTTCATTTTTCACGTTGTTGTTGTTGAACAC 660
Qy 660 GGTACTTGAACCTTATTAATAATTAATAATTTTATAGTTTCACTTTTATATAATAC 719
Db 661 GGTACTTGAACCTTATTAATAATTAATAATTTTATAGTTTCACTTTTATATAATAC 720
Qy 720 ----- 719
Db 721 TTCATTGATGCAATTTATATTTTGTGATGAATAATAAGAGTTTGTAAAAATATATTAT 780
Qy 720 TCATATAATATATAGGTTTAGAATGCCAATTTTAAAAAAGAAATAAAAAATAAATAG 779
Db 781 TCATATAATATATAGGTTTAGAATGCCAATTTTAAAAAAGAAATAAAAAATAAATAG 840
Qy 780 AATAAAATCGAAAAATGAATGTAAAAATTTGAGGGGGAACAATAAATAATGAAGTC 839
Db 841 AATAAAATCGAAAAATGAATGTAAAAATTTGAGGGGGAACAATAAATAATGAAGTC 900
Qy 840 TATTATTTAAATTTTCCATTAGAATTTCTATTTTCTTAGTTAAATGAGCTAGCCAGTTG 899
Db 901 TATTATTTAAATTTTCCATTAGAATTTCTATTTTCTTAGTTAAATGAGCTAGCCAGTTG 960
Qy 900 GGAGTACACGAAATGTGATGAACAGTTGCATGAGGAAATTAATGTAGTAGAGGGA 959
Db 961 GGAGTACACGAAATGTGATGAACAGTTGCATGAGGAAATTAATGTAGTAGAGGGA 1020
Qy 960 TAGCAGACAAAAATCCAGCCCAAGCTAGCTCTCACGCGAATCGATCCACAGTCCTT 1019
Db 1021 TAGCAGACAAAAATCCAGCCCAAGCTAGCTCTCACGCGAATCGATCCACAGTCCTT 1080
Qy 1020 TACAGAGTTTCAACCGGATGAATTCGATGCGCATGCACTGAACATTTGTTCTCAGCTG 1079
Db 1081 TACAGAGTTTCAACCGGATGAATTCGATGCGCATGCACTGAACATTTGTTCTCAGCTG 1140
Qy 1080 CCAAGTACCCCTCACACTCAACCAACCCCTTGTGTTTCTCCCATGATGATTAAGTCAAG 1139
Db 1141 CCAAGTACCCCTCACACTCAACCAACCCCTTGTGTTTCTCCCATGATGATTAAGTCAAG 1200

Qy 1140 TTTATCTCTTCTTCTGTTCTGGAATTTTCAAGGCTTCAACACGTCGAGCTCAATCTT 1199
Db 1201 TTTATCTCTTCTTCTGTTCTGGAATTTTCAAGGCTTCAACACGTCGAGCTCAATCTT 1260
Qy 1200 GTACCAAAACGCGCCCAAAAAAGAAATCTCATCCGTTCACTTTAGCCACTTAAAG 1259
Db 1261 GTACCAAAACGCGCCCAAAAAAGAAATCTCATCCGTTCACTTTAGCCACTTAAAG 1320
Qy 1260 CTAGCCAAACGCGTGTATCTTCTCTATATATTTAGTCTCTTAACACAAACACTACCA 1319
Db 1321 CTAGCCAAACGCGTGTATCTTCTCTATATATTTAGTCTCTTAACACAAACACTACCA 1380
Qy 1320 TTTATCAATTTCAAAACCTTGTCTTACTTACTACACACTAGAGAATA 1367
Db 1381 TTTATCAATTTCAAAACCTTGTCTTACTTACTACACACTAGAGAATA 1428

RESULT 3
SOYSEB7S
LOCUS Glycine max gene for basic 7S globulin, complete cds.
DEFINITION Glycine max gene for basic 7S globulin, complete cds.
ACCESSION D16107
VERSION D16107.1 GI:434060
KEYWORDS Glycine max (soybean)
SOURCE Glycine max
ORGANISM Glycine max
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.

REFERENCE 1
AUTHORS Kagawa,H. and Hirano,H.
TITLE Sequence of a cDNA encoding soybean basic 7S globulin
JOURNAL Nucleic Acids Res. 17 (21), 8868 (1989)
MEDLINE 90067863
PubMed 2587227

REFERENCE 2
AUTHORS Watanabe,Y. and Hirano,H.
TITLE Nucleotide sequence of the basic 7S globulin gene from soybean
JOURNAL Plant Physiol. 105 (3), 1019-1020 (1994)
MEDLINE 94336768
PubMed 8058830

REFERENCE 3 (bases 1 to 3393)
AUTHORS Watanabe,Y.
TITLE Cloning and sequence analysis of the basic 7S globulin gene in
soybean

JOURNAL Unpublished
REFERENCE 4 (bases 1 to 3393)
AUTHORS Watanabe,Y.
TITLE Direct Submission
JOURNAL Submitted (23-APR-1993) Yoshihiro Watanabe, National Institute of
Agrobiological Resources, Department of Molecular Biology; 2-1-2
Kannondai, Tsukuba, Ibaraki 305, Japan (Tel:81-298-38-7031,
Fax:81-298-38-7032)

FEATURES
source Location/Qualifiers
1..3393
/organism="Glycine max"
/mol_type="genomic DNA"
/cultivar="Miyagishirome"
/db_xref="taxon:3847"
/clone="pBGW1 and pBGW2"
/tissue_type="leaf, etiolated"
/clone_lib="pSOY-LIB"
/dev_stage="seedling"
1216..1230
/standard_name="heat shock like element"
/notes="HSE"
/evidence=not_experimental
1254..1257
TATA_signal
1344..1350
misc_feature
1369
/note="primary transcription start site"

| TITLE | Sequence and Analysis of Chromosome 2 of Dictyostelium |
|--|---|
| JOURNAL | Unpublished |
| REMARK | The Dictyostelium Genome Sequencing Consortium |
| REFERENCE | 2 (bases 1 to 56152) |
| AUTHORS | Baumgart,C. |
| TITLE | Direct Submission |
| JOURNAL | Submitted (04-APR-2002) Genome Analysis, Institute of Molecular Biotechnology, Beutenbergstr. 11, Jena 07745, Germany |
| COMMENT | CDS predictions from GenBank may contain errors. Further information is available from IMB Jena, Department of Genome Analysis (http://genome.imb-jena.de/dictyostelium/) and the University Cologne, Institute for Biochemistry I (http://www.uni-koeln.de/dictyostelium/project.shtml) Funding |
| FEATURES | <p>Agency : Deutsche Forschungsgemeinschaft (DFG).</p> <p>* NOTE: This is a 'working draft' sequence.</p> <p>* This sequence will be replaced</p> <p>* by the finished sequence as soon as it is available and</p> <p>* the accession number will be preserved.</p> <p>Location/Qualifiers</p> <p>1. .56152</p> <p>/organism="Dictyostelium discoideum"</p> <p>/mol_type="genomic DNA"</p> <p>/strain="AX4"</p> <p>/db_xref="taxon:44689"</p> <p>/chromosomes="2"</p> <p>/map="4790445-4846595"</p> |
| ORIGIN | |
| Query Match | 6.0%; Score 82.4; DB 2; Length 56152; |
| Best Local Similarity | 44.3%; Pred. No. 0.0011; |
| Matches 430; Conservative 0; Mismatches 531; Indels 9; Gaps 2; | |
| QY 13 | TACATTAAAAAATTTGGATTTTAAATTTTAGAAATTCATGATTTTATTTTATTTTATTTTACCAG 72 |
| Db 36134 | TACAAGATGAATTTGTTTATTGCAAATGAATTAATAAATAATAATCAACAACAACAAC 36075 |
| QY 73 | AAATCGGTTAATTTGTTTAAATATTACCACTAATTTATAAATTTTATTTTAGGCAAT 132 |
| Db 36074 | AAATGTTTGAATTAATCATCATTTTCAAAATTTTGTATATTGTATAAAAAAATAA 36015 |
| QY 133 | TAAGCATGTTTGTAAAAATATATATTTGTTATAATACTTTTCAAAAGTATAAAGTTGA 192 |
| Db 36014 | AAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 35955 |
| QY 193 | TGATGCGGTGGTGATGATTTTATTTAGTTCTAGGTCGAATCGCAAGTTGGTTAGACATT 252 |
| Db 35954 | TTTCTCTCCACTCTTTATTTAT 35895 |
| QY 253 | TAGCCTTATTTCTTTTCTCAACCAAAATAAATGTAATGGAACAACTTAGGAAAAAAA 312 |
| Db 35894 | ATATATATTATTATGCATATTAAAAAATAAATAAATAAATAAATAAATAAATAAATAA 35835 |
| QY 313 | GAAATCAAAATTTGAAACATCATCCGGTGGATCGGAAGCCCAACCAACCGTGACCCAA 372 |
| Db 35834 | AAAAACAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 35775 |
| QY 373 | CAATATTAAAAAAGAGTTTGCTCTACAGTAATGCGATCTTTTATTTTATCAATACTTTT 432 |
| Db 35774 | TAAAGTTGGTTTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTAT 35715 |
| QY 433 | TCCACTTCTAAATCTTGAGATTTGACCGGTTTAACATAATTAAGTTGTTATATCCACGGT 492 |
| Db 35714 | TTTATTATTATAATTTTATTATTATTATTATTATTATTATTATTATTATTATTATTATA 35655 |
| QY 493 | CCCTAAAAAAACTTGTGTACCGTGCCCTCACATTTCAACTTTTGCGCACCTTGAAAGCGGTA 552 |
| Db 35654 | ATAAATGTAAGTAAAAATATTATCAATAATTTTTTTTTTTTTTAAAAAATAAAGTGGATT 35595 |
| QY 553 | TGTTTAGGTTAGTGTGTTGCAACAGTTGAAGGCGATCACTCAGGAGGCTACTTGGTCTTGC 612 |
| Db 35594 | AAATCTAGGTTTCTATTATACAAAAAATAACCAATTATATAATTTTTTTTTTTTTTTTT 35533 |

| | | |
|----|---------------------------|---|
| | Query Match | 6.0%; Score 82.4; DB 3; Length 183648; |
| | Best Local Similarity | 44.3%; Pred. No.: 0.00089; |
| | Matches 430; Conservative | 0; Mismatches 531; Indels 9; Gaps 2; |
| QY | 13 | TACATTAAAAAATTTTGATTTTTAAAATTTTAGAAATTCATGATTTTATTTTATTTTACCAG 72 |
| DB | 34352 | TACAAGATGAATTTGTGTTATTGCCAAATGAATTAATAATAATAATCAACAACAACA 34411 |
| QY | 73 | AATCCGGTAAATCTGTTAAATAATTACCAACTAATTATTAATTAATTTATTTTAAGCGCAT 132 |
| DB | 34412 | AAATCTGTTGAAATTAATCATCATTTTCAAATTTTGTATATATTTGTATAAAAAAAAA 34471 |
| QY | 133 | TAAAGCATGTTTGTAAAAATATATATATGTTATAATACTTTCCAAAGTATAAAGTTGA 192 |
| DB | 34472 | AATCCTTGTCATCTT 34531 |
| QY | 193 | TGATGGGTGGTGTAGATTAITTTTAGTTC TAGGTTCGAATCGCAAGTTGGTTTAGACATT 252 |
| DB | 34532 | TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTCTCTCCATCTCTTTATTTTAT 34591 |
| QY | 253 | TAGCCTTATCTTTTTTCTAACCAAAATAAATGTAAATGGAAGAAACCTTTAGGAAAAAAA 312 |
| DB | 34592 | ATATATATTTATTATGCATATTAAAAAATAAAAAAAAAATAAAAAATAATAATAATAA 34651 |
| QY | 313 | GAAATCAAAATGAAAACATCATCCGGTGGAGTCGAGAAGCCCCCACCCACGTGACCCA 372 |
| DB | 34652 | AAAAAATAAAAAAAAAAAAAAAAAAATGTAATAATAAANAACAGAGAAATTTTAAAACT 34711 |
| QY | 373 | CAATATTAATAAAGAGTTTGCTCTCAGATAATGCGATACCTTTTTTATTCATACCTTT 432 |
| DB | 34712 | TAAAGGTGGTTTTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTA 34771 |
| QY | 433 | TCCACTTCAAAATCTTGGNATTTGCACGGTTAACTAATTAAGTGTATATCCACGGT 492 |
| DB | 34772 | TTTATTTATTAATAATTTTATTTTATTTTATTTTATTTTATTTTATTTTAAAGGAACA 34831 |
| QY | 493 | CCTAAAAAAACTTGTGTACCGTGCCTTCACATTTCAACTTTGCGCACCCCTGAAGCGTTA 552 |
| DB | 34832 | ATAAATGAACTGAAAATATTTCATCAAAATTTTTTTTTTAAAAAATAAAGTGGATTTT 34891 |
| QY | 553 | TGTTTAGGTAGTGTGTTTGCACAGTGTGAAGCGCATCACTCAGAGGCGTACTTGTCCTGC 612 |
| DB | 34892 | AAATCTAGGTTTCTATTATCACAAAAATAATACCAATTTATAAATTTTTTTTTTTTTT 34951 |
| QY | 613 | TTTTGGCTCTTTGTTCAATTTTTTCAG - - - - TGATTTTGTGGTCAACACGGGTAC 665 |
| DB | 34952 | TTTTTTTTTTTTTTTTTTTTTTTTTTTATTATTTATTTATTTTATTTTATTTTATTTAT 35011 |
| QY | 666 | TGAAACTTATATAAATTAACAFAATTTTATAAGTTTCCTCTTTATAATAAATCTCATAT 725 |
| DB | 35012 | TTTTTTTTTTTTTTTATTTCTTTTTTTTTTTCTTTCTGATTTCTTTTTTTCGCAAAAAC 35071 |
| QY | 726 | AATATATPAGGTTTATAGATG - CCAATTTTTTAAAAAAGAAATAAAAAATAATAAGAA 783 |
| DB | 35072 | AATAATAAATTAATAAAAAAAAAAATAATTAATAAATAAATAAATAAATAAATAAATAA 35131 |
| QY | 784 | AAATCGAAAAAATGAAATGTAAAAAATTTGAGGGGACAAATAAAATATGAAAGTCTATT 843 |
| DB | 35132 | AAAAAAAAAAAAAAAAAATAATTAATAATTAATAAATAAGTAAATAATAATAATAATAATA 35191 |

QY 844 ATTTAAATTTTCATTAGAAATTCATTTTCCCTAGTTAATGACGTAGCCAGTGGGAG 903

Db 35192 TTAACACGAGATTAAAGAATTCATCCCTAAAAAGAAAAATAAAAAATAAAAAAATA 35251

QY 904 ATACACGAAATGTCATGAAACAGTTGCATGTAGGGAATTAATCTAGTAGGAGGATAGC 963

Db 35252 AAATAAAAAAATAACACGACAAATAATTAAACAAAAAATAAAAAAAGATTACC 35311

QY 964 AAGACAAAAA 973

Db 35312 AGTGTAAAAA 35321

RESULT 6
AC084128/c
LOCUS
DEFINITION Homo sapiens chromosome 8, clone CTD-3118D11, complete sequence.
AC084128
VERSION AC084128.6 GI:22004505
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 129240)
Birren,B., Nusbaum,C. and Lander,E.
Homo sapiens chromosome 8, clone CTD-3118D11
JOURNAL
TITLE
2 (bases 1 to 129240)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Beda,F., Boguslavsky,L.,
Bouckigalter,B., Brown,A., Burkett,G., Campopiano,A., Castle,A.,
Choepel,Y., Colangelo,M., Collins,S., Collamore,A., Cooke,P.,
DeArallano,K., Dewar,K., Diaz,J.S., Dodge,S., Ferreira,P.,
FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Goyette,M.,
Graham,L., Grand-Pierre,N., Hagos,B., Heaford,A., Horton,L.,
Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A., LaRocque,K.,
Lamarez,R., Landers,T., Lehoczy,J., Levine,R., Lieu,C., Liu,G.,
Macdonald,P., Marquis,N., McCarthy,M., McEwan,P., McKernan,K.,
McPheters,R., Meldrim,J., Meneus,L., Mihova,T., Mlenga,P.,
Morrow,J., Murphy,T., Naylor,J., Norman,C.H., O'Connor,T.,
O'Donnell,P., O'Neil,D., Oliver,T.M., Olvera,J., Peterson,K.,
Pierre,N., Pisani,C., Pollara,V., Raymond,G., Rieback,M., Riley,R.,
Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S., Severy,P.,
Sougnuez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Tirrell,A., Travers,M., Triglio,J., Vassiliev,H., Viel,R., Vo,A.,
Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J.,
Zimmer,A. and Zody,M.
Direct Submission
Submitted (13-OCT-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE
3 (bases 1 to 129240)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L.,
Bouckigalter,B., Brown,A., Camarato,J., Campopiano,A., Chang,J.,
Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collamore,A.,
Cooke,A., Cooke,P., DeArallano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Farlo,S., Ferreira,P., FitzGerald,M., FitzHugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Gird,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I.,
Johnson,R., Jones,C., Kanat,A., Karatas,A., Kellis,C., LaRocque,K.,
Lamarez,R., Landers,T., Lehoczy,J., Levine,R., Lindblad-Toh,K.,
Liu,G., Maclean,C., Macdonald,P., Major,J., Marquis,N.,
Matthews,C., McCarthy,M., McEwan,P., McKernan,K., Meldrim,J.,
Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neil,D., Orbu,C., Peterson,K., Phunkhang,P., Pierre,N.,
Pollara,V., Raymond,G., Retta,R., Rieback,M., Riley,R., Rise,C.,
Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S.,
Schupack,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S.,
Theodore,J., Topham,K., Travers,M., Travis,N., Triglio,J.,

Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
 Direct Submission
 Submitted (02-JUN-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 4 (bases 1 to 129240)
 Biren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Bana, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhgalter, B., Canarata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A., Cook, A., Cooke, P., DeAveliano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliiev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., MacLean, C., Macdonald, P., Major, J., Matthews, C., McCarthy, M., Meldrim, J., Menus, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schuback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE
JOURNAL
COMMENT
 Submitted (30-JUL-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Jul 30, 2002 this sequence version replaced gi:21306731.
 All repeats were identified using RepeatMasker:
 Smit, A. F. A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center -----
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WBIR
 Web site: http://www.seq.wi.mit.edu
 Contact: sequence.submissions@genome.wi.mit.edu
 ----- Project Information -----
 Center project name: L11383
 Center clone name: 3118_D_11

FEATURES
 source
 1..129240
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="8"
 /map="8"
 /clone="CTD-3118D11"
 /clone_lib="CITD2 Human BAC"
 1..224
 /rpt_family="AluSg/x"
 225..403
 /rpt_family="AluUb"
 669..855
 /rpt_family="L1MC3"
 1519..1600
 /rpt_family="AT rich"
 complement(1619..1895)
 /rpt_family="AluSc"
 complement(1897..2197)
 2310..2593
 /rpt_family="AluSg"
 complement(2598..3451)
 /rpt_family="L1MA4A"
 2966..2969
 /note="single clone coverage"
 3483..3713
 /rpt_family="L1MDa"
 complement(3714..4017)
 /rpt_family="AluSg"
 3739..3748
 /note="<30 qual SNGL region"

unsure
 3785..3811
 /note="<30 qual SNGL region"
 4018..4255
 /rpt_family="L1MDa"
 4267..4982
 /rpt_family="L1MDa"
 4974..5464
 /rpt_family="L1MD2"
 5827..6015
 /rpt_family="MIR"
 8136..8345
 /rpt_family="MIR"
 8552..8737
 /rpt_family="L2"
 9347..9644
 /rpt_family="AluSx"
 complement(10188..10324)
 /rpt_family="MIR3"
 10770..11002
 /rpt_family="MIR"
 complement(11116..11638)
 /rpt_family="MER1A"
 12153..12222
 /rpt_family="L2"
 complement(12668..12878)
 /rpt_family="MIR"
 complement(13676..14221)
 /rpt_family="MER41A"
 complement(15056..15245)
 /rpt_family="AluUb"
 15272..15714
 /rpt_family="AluSx"
 complement(16162..16386)
 /rpt_family="MIR"
 16712..17008
 /rpt_family="AluJo"
 complement(18962..18963)
 /rpt_family="MERSA"
 complement(19028..19089)
 /rpt_family="L2"
 19090..19375
 /rpt_family="AluUb"
 complement(19376..19430)
 /rpt_family="L2"
 complement(19480..19643)
 /rpt_family="L2"
 complement(19744..19829)
 /rpt_family="MIR3"
 complement(19976..20109)
 /rpt_family="MIR"
 22048..22074
 /rpt_family="(T)n"
 complement(22240..22864)
 /rpt_family="L1PA7"
 complement(23650..23959)
 /rpt_family="AluUb"
 complement(24055..24374)
 /rpt_family="AluSg1"
 24894..24922
 /rpt_family="AT rich"
 complement(24934..25223)
 /rpt_family="AluUb"
 26058..26354
 /rpt_family="AluSx"
 26936..27229
 /rpt_family="AluSx"
 28099..28243
 /rpt_family="AluSx"
 29591..29611

Query Match
 Best Local Similarity 6.0%; Score 81.6; DB 9; Length 129240;
 Matches 393; Conservative 0; Mismatches 484; Indels 1; Gaps 1;

Carucci,D.J., Hoffman,S.L., Newbold,C., Davis,R.W., Fraser,C.M. and
Barrell,B.
TITLE Genome sequence of the human malaria parasite Plasmodium falciparum
JOURNAL Nature 419, 498-511 (2002)
REFERENCE 3 (bases 1 to 12029)
AUTHORS Gardner,M.J.
TITLE Direct Submission
JOURNAL Submitted (02-OCT-2002) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA
FEATURES
source
1..12029
/organism="Plasmodium falciparum 3D7"
/mol_type="Genomic DNA"
/isolate="3D7"
/db_xref="taxon:36329"
/chromosome="2"
complement(1..53)
/note="AT rich"
/rpt_type=tandem
58..97
/rpt_type=tandem
complement(58..97)
/rpt_type=tandem
/rpt_unit="(TA)n"
complement(117..156)
/note="AT rich"
/rpt_type=tandem
153..204
/rpt_type=tandem
complement(157..204)
/rpt_type=tandem
/rpt_unit="(TA)n"
complement(241..264)
/note="AT rich"
/rpt_type=tandem
complement(335..356)
/rpt_type=tandem
/rpt_unit="(A)n"
complement(372..513)
/note="AT rich"
/rpt_type=tandem
557..589
/rpt_type=tandem
complement(557..589)
/rpt_type=tandem
/rpt_unit="(TA)n"
complement(659..684)
/note="AT rich"
/rpt_type=tandem
709..844
/rpt_type=tandem
/rpt_unit="(TAAA)n"
complement(916..961)
/rpt_type=tandem
/rpt_unit="(TAAAA)n"
complement(1035..1061)
/note="AT rich"
/rpt_type=tandem
complement(1076..1122)
/rpt_type=tandem
/rpt_unit="(TA)n"
complement(1284..1317)
/rpt_type=tandem
/rpt_unit="(TA)n"
1284..1313
/rpt_type=tandem
complement(1374..1672)
/note="AT rich"
/rpt_type=tandem
complement(1722..1810)
/note="AT rich"
/rpt_type=tandem
1931..1987

repeat_region
complement(1931..1967)
/rpt_type=tandem
/rpt_unit="(TA)n"
repeat_region
complement(1968..1999)
/note="AT rich"
/rpt_type=tandem
complement(2170..2229)
/note="AT rich"
/rpt_type=tandem
2381..2414
/rpt_type=tandem
/rpt_unit="(A)n"
2384..2414
/rpt_type=tandem
complement(2417..2460)
/note="AT rich"
/rpt_type=tandem
complement(2545..2566)
/note="AT rich"
/rpt_type=tandem
complement(2615..2643)
/note="AT rich"
/rpt_type=tandem
complement(2840..2869)
/note="AT rich"
/rpt_type=tandem
complement(3000..3045)
/rpt_type=tandem
/rpt_unit="(TA)n"
3010..3045
/rpt_type=tandem
complement(3049..3086)
/rpt_type=tandem
/rpt_unit="(TAA)n"
3049..3082
/rpt_type=tandem
3282..3324
/rpt_type=tandem
/rpt_unit="(A)n"
complement(3325..3419)
/rpt_type=tandem
/rpt_unit="(TAAA)n"
complement(3461..3483)
/rpt_type=tandem
/rpt_unit="(TA)n"
complement(3510..3691)
/note="AT rich"
/rpt_type=tandem
3705..3745
/rpt_type=tandem
/rpt_unit="(A)n"
complement(3766..3797)
/rpt_type=tandem
/rpt_unit="(A)n"
complement(3798..3935)
/note="AT rich"
/rpt_type=tandem
complement(4077..4142)
/note="AT rich"
/rpt_type=tandem
complement(4189..4236)
/rpt_type=tandem
/rpt_unit="(TA)n"
4201..4236
/rpt_type=tandem
complement(4265..4348)
/rpt_type=tandem
/rpt_unit="(TATAA)n"
complement(4422..4592)
/note="AT rich"
/rpt_type=tandem
4593..4618

| | |
|-----------|--|
| AUTHORS | |
| TITLE | |
| JOURNAL | |
| REFERENCE | |
| AUTHORS | |
| TITLE | |
| JOURNAL | |
| REFERENCE | |
| AUTHORS | |
| TITLE | |
| JOURNAL | |
| COMMENT | |
| FEATURES | |
| source | |
| CDS | |
| CDS | |
| CDS | |

artificial sequences.

REFERENCE

1. Berlin, K., Braun, A., Dietler, J., Gueig, D., Howe, A., Mueller, J., Olek, A., Piepenbrock, C., Adorjan, P., Grabs, G., Lesche, R., Leu, E., Lewin, A., Lipscher, E., Maier, S., Model, F., Mueller, V., Otto, T., Pelet, C. and Ziebarth, H.
METHODS AND NUCLEIC ACIDS FOR THE ANALYSIS OF HEMATOPOIETIC CELL PROLIFERATIVE DISORDERS
Patent: WO 02077272-A 386 03-OCT-2002;
EpiGenomics AG (DE)

FEATURES

source
Location/Qualifiers
1..8056
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:38630"
/note="chemically treated genomic DNA (Homo sapiens)"

ORIGIN

Query Match 5.5%; Score 74.6; DB 6; Length 8056;
Best Local Similarity 44.6%; Pred. No. 0.025;
Matches 393; Conservative 0; Mismatches 474; Indels 14; Gaps 2;

QY 9 TCTATACATTAATAAATTGATTTTAAATTTTGAATTCATGATTTTATTTTTTTTAA 68
Db 1833 TTTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1892
QY 69 CCAGAAATCCGTTAATATGTTTAAATATTACCACTAATTTATATAATTTTATTTAAGG 128
Db 1893 TTATATAAATAAATAAATTATATAAATAAATAAATAAATAAATAAATAAATAAATAA 1951
QY 129 CAATTAAGCATGTTTGAATAAATAATATATATGTTTATAAAT-----ACTTTT 175
Db 1952 AATATAAATTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATT 2011
QY 176 CAAAAAGTATAAGTTGATCATCGCTGGTGGTAGATTTATTTAGTCTAGGTCGAATGC 235
Db 2012 TAAATATATTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 2071
QY 236 AAGTTGGTTAGACATTTAGCCTTATCTTTTCTAACCAAAATAAATGTAATGGAAA 295
Db 2072 AAAATATTTTAAATATTATAAATAAATAAATAAATAAATAAATAAATAAATAAATT 2131
QY 296 ACCTTTAGAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 355
Db 2132 TTTATTTTATTAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATT 2191
QY 356 ACACCCACGTGACCCCAACAATAATTAAAAATAAGAGTTTCTCTACAGTAATGCGATCTT 415
Db 2192 AAAAAATGAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATT 2251
QY 416 TTTTATTCATATCTTTTCCACTTCTAAAACTTGGAGATTGGACCGTTAACTAATAATA 475
Db 2252 TAATATTTAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAAT 2311
QY 476 GTGTTATATCCACCGTCTTAAAAAATACTGTGTACCGTCTCACATTTCAACTTTGCG 535
Db 2312 TTATTAATTTTGTATTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATT 2371
QY 536 CACCTGAAAGCCGTTATGTTAGTGTAGTTGTTGCAACAGTTGGAAGCGCATCACTCAGG 595
Db 2372 ATAAATAATTTATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATT 2431
QY 596 AGGCTACTGTGCTGCTTTGCGCTTTTGGCTTTTGTCAATTTTTCAGTGAATTTGTGTGA 655
Db 2432 TTTTAAAAATAAATTTTATTTTATTTTATTTATATATATTTTATTTATTTTGTGTTAAAA 2491
QY 656 ACACGCGTACTGAAACCTTATTAAATAATACATAAATTTTATAAGTTTTCACCTCTATATA 715
Db 2492 TAAATTAATAATTAATGAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 2551
QY 716 ATACTCATATATATATAGGTTTGAATGCAATTTTAAAAAATAAATAAATAAATAAATAA 775
Db 2552 ATTTAAAAATTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 2611

RESULT 13

AC115598/LOCUS 154071 bp DNA linear INV 12-MAR-2003
DEFINITION Dictyostelium discoideum chromosome 2 map 581427-735498 strain AX4, complete sequence.
AC115598 AC115593 AC117269
VERSION AC115598.2 GI:28850406
KEYWORDS HTG.
SOURCE Dictyostelium discoideum
ORGANISM Dictyostelium discoideum
Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
REFERENCE 1 (bases 1 to 154071)
AUTHORS Gloeckner, G., Eichinger, L., Szafarski, K., Pachebat, J., Dear, P., Lehmann, R., Baumgart, C., Parra, G., April, J.F., Guigo, R., Kumpf, K., Tunggal, B., Cox, E., Quail, M.A., Platzner, M., Rosenthal, A. and Noegel, A.A.
Sequence and analysis of chromosome 2 of Dictyostelium discoideum
Nature 418 (6893), 79-85 (2002)
12097910
22092622
The Dictyostelium Genome Sequencing Consortium
REFERENCE 2 (bases 1 to 154071)
AUTHORS Baumgart, C.
Direct Submission
JOURNAL Submitted (21-MAR-2002) Genome Analysis, Institute of Molecular Biotechnology, Beutenbergstr. 11, Jena 07745, Germany
REFERENCE 3 (bases 1 to 154071)
AUTHORS Baumgart, C.
Direct Submission
JOURNAL Submitted (05-MAR-2003) Genome Analysis, Institute of Molecular Biotechnology, Beutenbergstr. 11, Jena 07745, Germany
REFERENCE 4 (bases 1 to 154071)
AUTHORS Baumgart, C.
Direct Submission
JOURNAL Submitted (12-MAR-2003) Genome Analysis, Institute of Molecular Biotechnology, Beutenbergstr. 11, Jena 07745, Germany
COMMENT On or before Mar 5, 2003 this sequence version replaced GI:19569986, GI:20087116, GI:19570017.
CDS predictions from GeneID do not necessarily reflect true genes. Further information is available from IMB Jena, Department of Genome Analysis
(http://genome.imb-jena.de/dictyostelium/)
and the University Cologne, Institute for Biochemistry I
(http://www.uni-koeln.de/dictyostelium/project.shtml)

FEATURES

source
Location/Qualifiers
1..154071
/organism="Dictyostelium discoideum"
/mol_type="genomic DNA"
/strain="AX4"
/db_xref="taxon:44689"
/chromosomes="2"
/map="581427-735498"
/notes="GeneID exon scores (in order of location ranges): 110.23, 208.32, 33.09 - GSCJ_ID dd_03170"
/codon_start=1
/product="similar to Plasmodium falciparum. Hypothetical protein"
/protein_id="AA053175.1"
/db_xref="GI:28850407"
/translation="MKFIPNLYDNQNLPSNDEDDPIDVIASLNDFLMSTKNGEILH"
Agency : Deutsche Forschungsgemeinschaft (DFG).
Funding
Location/Qualifiers
1..154071
/organism="Dictyostelium discoideum"
/mol_type="genomic DNA"
/strain="AX4"
/db_xref="taxon:44689"
/chromosomes="2"
/map="581427-735498"
/notes="GeneID exon scores (in order of location ranges): 110.23, 208.32, 33.09 - GSCJ_ID dd_03170"
/codon_start=1
/product="similar to Plasmodium falciparum. Hypothetical protein"
/protein_id="AA053175.1"
/db_xref="GI:28850407"
/translation="MKFIPNLYDNQNLPSNDEDDPIDVIASLNDFLMSTKNGEILH"
CDS

| | | | |
|----|--------|--|---------|
| Db | 105064 | AAAAAAAAAAGAAATTTAATAAAAAATAGAAATTTAAAAATAGAAAAATTAATAATAGAAATTT | 1050005 |
| Qy | 62 | TTTTTTACAGAAATCCGTTAAATATTGTTAAAAATATTACCAACTAATTTATAAAATTTTAT | 121 |
| Db | 105004 | AAACTTACTTTATAAGGATTAATATTAAACAATTTTAAACTTTATATATACTTCTCAG | 104945 |
| Qy | 122 | TTTAAGGCAATTAAGCATGTTTGATAAAATATATATATTGTTATAAATACTTTTCAAAG | 181 |
| Db | 104944 | TTATTTACACATAGAAATAGTTTAAATTTCTTTTTATTTTGAACACTATATTTTTTTTTTTT | 104885 |
| Qy | 182 | TATAAAGTTGATGATGGCGTGGTGTAGATTAATTTAGTCTTAGGTTGCAATGCAAGTTG | 241 |
| Db | 104884 | TCCTTTTTTCTTTTTCTTTTGTGTAAAAATAAATAAATAAATAAATAAATAAATAAATAA | 104825 |
| Qy | 242 | GTTTACACATTTAGCGTTATCTCTTTTCTACCAAAATAAATGTAAATGGAAGAACCTTT | 301 |
| Db | 104824 | TTTTTTTTTTTTTTTTTTTTTTTTTTTGTGATGATTTAAATTTACTGTTCTATATTTAACTTT | 104765 |
| Qy | 302 | AGGAAAAAAGAAATCAAAATTTGAAATCATCATCCGGTGGAGTCGAGAGCCACACC | 361 |
| Db | 104764 | GGAAAAAGAAAAAATAAATAATTTAGAAATTTAAAAATAAANAATTTTAAATTTAAATTT | 104708 |
| Qy | 362 | AGTGACCCACAATATTAAANAATAGAGTTTGCTCTACAGTAAATCGGATACTTTTTTAT | 421 |
| Db | 104707 | ATAAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA | 104648 |
| Qy | 422 | TCAATACTTTTTCCACTCTTAAATCTTGGAGATTTGCACCGTTAACTTAATTTAAAGTGTA | 481 |
| Db | 104647 | AAAAAATAAATAATGATTTTTGTGAATCGAATTTTATTCACAAATCATGATTACCCATTA | 104588 |
| Qy | 482 | TATCCACGGTCTTAAAAAACTTGTGACCGTGCTCACATTTCAACTTTTCGCGACCGT | 541 |
| Db | 104587 | TTTAGATTTTATAAAAAATAAATAATTTTAAATTTTACCTTAAAAATAAATTTTTTAT | 104528 |
| Qy | 542 | GAAAGCCGTTATGTTAGTTAGTGTGTTGCAACAGTTGAAGCGCATCACTCAGGAGGCTA | 601 |
| Db | 104527 | TAAAAATAAATAAATAATTAATTAATTAATTAATAAATTAATAAATCTATATTAAATATTATTA | 104468 |
| Qy | 602 | CTTGGTCTGCTTTTGCGTCTTTGTTTCAATTTTTTCAGCTGATTTTGTGTGGAACACGC | 661 |
| Db | 104467 | TTATAATTCCGTGTGTGTTA | 104408 |
| Qy | 662 | GTACTTGAACCTTATATAAATACAT--AATTTATAAGTTTCACTCTTCTTATATAAATAC | 719 |
| Db | 104407 | CAATTTTATATTTTGAATAAGTAATAAATAAATAAATAAATAAATAAATAAATAAATAA | 104348 |
| Qy | 720 | TCATATAATATATAGGTTTAGAATGCCAATTTTTTAAAAAAGAAATAAAAAATAAATAG | 779 |
| Db | 104347 | AAATAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA | 104288 |
| Qy | 780 | AATAAAAATCGAAAAAATGAANAATGTAAAAAATTTGAGGGGACAAATAAANAATATGAAGTC | 839 |
| Db | 104287 | TTAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA | 104228 |
| Qy | 840 | TATTAATTTAAATTTTCCATTAAGAAATCTATTTTTCTTTAGTTTAAATATGACGTAGCCAGTG | 899 |
| Db | 104227 | TTTTCCAACTATTTTAAATTTAAATTTTTTTTATGTTTTTTTATTATTAATCAATTTGTTTT | 104168 |
| Qy | 900 | GGAGATACAGAAATGTCTGAACA | 926 |
| Db | 104167 | GTGATATCATTTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA | 104141 |

RESULT 14
AC116305_3
WPCOMMENT

| Sequence | split into | 5 fragments | LOCUS | AC116305 | Accession | AC116305 |
|---------------|------------|-------------|-------|----------|-----------|----------|
| Fragment Name | Begin | End | | | | |
| AC116305_0 | 1 | 110000 | | | | |
| AC116305_1 | 100001 | 210000 | | | | |
| AC116305_2 | 200001 | 310000 | | | | |
| AC116305_3 | 300001 | 410000 | | | | |
| AC116305_4 | 400001 | 413138 | | | | |

Continuation (4 of 5) of AC116305 from base 300001 (AC116305 Dictyostelium discoideum chr

| | Query Match | 5.4%; | Score 73.6; | DB 3; | Length 110000; |
|----|-----------------------|----------------|--|-----------------|-------------------|
| | Best Local Similarity | 45.3%; | Pred. No. 0.022; | | |
| | Matches 388; | Conservative | 0; | Mismatches 459; | Indels 9; Gaps 3; |
| QY | 7 | AATCTATACATTAA | AAAAATTTGATTTTAAAAATTTAGAAATTCATGATTTTATTTTTTTT | 66 | |
| Db | 51226 | AATTCATAATAAC | TTGACCAGGTTTAAATTTGCAGGTACAAAATAAACCTTCAAT | 51285 | |
| QY | 67 | TACAGAAATCCG | TAAATATGTTAAAAATATTACCAACTAAATTTATATAAATTTATTTTAA | 126 | |
| Db | 51286 | AATTTTGATTTT | GAATATATATGATGCAATACCTTTTTCAAATCTATAATCTGATATATTA | 51345 | |
| QY | 127 | GGCAATTAAGCA | TGTTTGATAAAATATATATTTGTTATAAAT---ACITTTTCARAAGT | 182 | |
| Db | 51346 | ATTTCAAATGGA | GTCCAAATATATTTCCAAATTAATTTAGTTTGGTATATTAAATAATA | 51405 | |
| QY | 183 | ATAAAGTTGATG | AGCGGTGGTGAGATTAATTTTAGTTCTAGGTCGAATGCAAGTTGG | 242 | |
| Db | 51406 | ATCTTATTATAA | TATGATTTGTAACATTTAAATTTGAATATTAAATGATACATTC | 51465 | |
| QY | 243 | TTTAGACATTT | TAGCCTTATCTTTTTCTTACCAAAATAAATGATAATGGAACCTTTTA | 302 | |
| Db | 51466 | ATTTCATCATAA | CACTTTTTTTTTTTTTTAAAAAATAAATAATTAATTAATAATTA | 51525 | |
| QY | 303 | GGAAAAAAGAAAT | CAAAATTTGAAAACATCATCCGGTGGAGTCGAGAGCCACACCCA | 362 | |
| Db | 51526 | ATAAATATAAT | TATAATTTATAATTAATTAATTAATTAATAAATAATGAATAATAAT | 51585 | |
| QY | 363 | CGTGACCCACA | CAATATTTAAATPAGAGTTTGTGC--TACAGTAAATGCGATACTTTTT | 419 | |
| Db | 51586 | AATTACCTTGAC | CAAAAAATATCAAGAGGATAACGAATAAATAATTTGAAGTTAAATTTG | 51645 | |
| QY | 420 | ATTCAAATACT | TTTTTCCACTTCTAATAATCTTGAGATTTTGCACGGTTTAACTAATTAAGTG | 479 | |
| Db | 51646 | ATTCAATATAAT | CAACCAATTTAATTTTTTTTGAATTTCCCTGCAAAATCAATAACAGATA | 51705 | |
| QY | 480 | TATATCCAACG | GTCTCTAAAAAACTTGTGTACCGTGCCTTCACATTTTCAACTTTTGGCACC | 539 | |
| Db | 51706 | TTGAAAAATCTG | ATGAATAAAATGGACAATCTAAAGTATCATGAACATAAAACCAITTTT | 51765 | |
| QY | 540 | CTGAAGCCGTAT | GTGTTAGGTAGTGTGTCACACAGTTGAAGCGCATCACTCAGGAGGC | 599 | |
| Db | 51766 | CAATATTATTT | CTCTTTTGAATTTAAAGATCTCAAACATAAAACCTTTTTTACAAATATAA | 51825 | |
| QY | 600 | TACTTGGCTTG | CTTTTCCGCTCTTTTGTTCAAITTTT--TCACGTGATTTTGTGGTGAAC | 657 | |
| Db | 51826 | TCATCTTAAT | ACCTTGTTGATTTCTGTGTGAATATAATCTGTGATGTTGTTAATGATA | 51885 | |
| QY | 658 | ACGGTACTTGA | AACTTATTATAAATTAACATAATTTTATAAGTTTTCACTTCTTATATAAT | 717 | |
| Db | 51886 | TTCCAACCTGT | GTGCCAAGTAAATAATCTCATATTATAATTTAAATTTTGGTGGGTATA | 51945 | |
| QY | 718 | ACTCATATATA | TATAGGGTTTGAATGCCAATTTTTTAAAAAAGNATAAAAAAATAAAT | 777 | |
| Db | 51946 | TATCATTTGCT | TATATATATAAAAAAATAAATAAATAAATAAATAAATAAATAAATAA | 52005 | |
| QY | 778 | AGATAAAATCG | AAAAAATGAAATGTAAAAAATTTTGAGGGGACAAATAAATAATCAAG | 837 | |
| Db | 52006 | ATAATTAAT | TAAAAATAATAATATAATAATTAATTTAAATTAATACATACAAAAAATTTG | 52065 | |
| QY | 838 | TCTATTAAT | TAAAAATTT 853 | | |
| Db | 52066 | TTTAAACTATG | TAAAT 52081 | | |

| | |
|-----------|------------|
| RESULT 15 | DEFINITION |
| AC096550 | ACCESSION |
| LOCUS | VERSION |

KEYWORDS SOURCE ORGANISM

HTG.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE AUTHORS

1 (bases 1 to 101034)
Sulston, J.E. and Waterston, R.
Toward a complete human genome sequence
Genome Res. 8 (11), 1097-1108 (1998)
99063792
PUBMED
9847074

REFERENCE AUTHORS

2 (bases 1 to 101034)
Dauphin, S. and Meyer, R.
The sequence of Homo sapiens BAC clone RP11-20F13
Unpublished (2001)

REFERENCE AUTHORS

3 (bases 1 to 101034)
Waterston, R.H.
Direct Submission
Submitted (18-SEP-2001) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
4 (bases 1 to 101034)
Waterston, R.

REFERENCE AUTHORS

Direct Submission
Submitted (01-MAR-2002) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Mar 1, 2002 this sequence version replaced gi:15638715.

COMMENT

Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu/gsc>
Contact: sapiens@watson.wustl.edu

Summary Statistics
Center project name: H_NH0020F13
Drafting Center: WIBR

NOTICE: This sequence may not represent the entire insert of this
clone. It may be shorter because we only sequence overlapping
clone sections once, or longer because we provide a small overlap
between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by sequence
from more than one subclone; and the assembly was confirmed by
restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D.
McPherson, Department of Genetics, Washington University, St. Louis
MO. For additional information about the map position of this
sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

The RP11-11 human BAC library was made from the blood of one male
donor, as described by Osogawa, K., Woon, P.Y., Zhao, B., Frengen, E.,
Tateno, M., Catanesi, J.J. and de Jong, P.J. (1998) An improved
approach for construction of bacterial artificial chromosome
libraries. Genomics 51:1-8. The clone may be obtained either from
Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong
and coworkers at <http://www.chori.org>
VECTOR: pBACE3.6

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the right is RP11-335B8, 2000 bp overlap.
Actual start of this clone is at base position 1 of RP11-20F13;
actual end is at base position 70665 of RP11-335B8.

Data from AC083850 was used to finish this clone, AC013666.

FEATURES Source

The sequence of AC013666 has been incorporated into AC096550.

Location/Qualifiers
1..101034
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="2"
/map="2"
/clone="RP11-20F13"
/clone_lib="RPC1-11"
1..1075
/rpt_family="L1"
1376..1837
/rpt_family="L1"
1887..2087
/rpt_family="L1"
1940..1982
/rpt_family="AT_rich"
2175..2209
/rpt_family="AT_rich"
2197..2409
/rpt_family="L1"
2425..2447
/rpt_family="AT_rich"
2429..3528
/rpt_family="L1"
3595..3882
/rpt_family="Alu"
4204..4285
/rpt_family="ERV1"
4240..4283
/rpt_family="(TG)n"
4588..4628
/rpt_family="(TG)n"
5235..5660
/rpt_family="L2"
5685..5928
/rpt_family="MIR"
6002..6091
/rpt_family="L2"
7654..7924
/rpt_family="MaLR"
8205..8279
/rpt_family="MaLR"
8324..9069
/rpt_family="L1"
11481..11501
/rpt_family="AT_rich"
12371..12573
/rpt_family="ERV1"
12674..12864
/rpt_family="ERV1"
14213..14333
/rpt_family="L1"
14487..15130
/rpt_family="ERV1"
15471..15876
/rpt_family="ERV"
17863..17889
/rpt_family="AT_rich"
17904..18466
/rpt_family="ERV1"
18122..18142
/rpt_family="AT_rich"
18495..18517
/rpt_family="AT_rich"
18528..18865
/rpt_family="MER2_type"
18866..18920
/rpt_family="AT_rich"
18903..19656
/rpt_family="L1"
20507..20646

```

repeat_region /rpt_family="MER1_type"
21297..21413 /rpt_family="ACHobo"
21864..22023 /rpt_family="MIR"
22505..22728 /rpt_family="MIR"
23335..23569 /rpt_family="MALR"
24524..24994 /rpt_family="MALR"
25131..25371 /rpt_family="L2"
26055..26246 /rpt_family="MIR"
27295..27899 /rpt_family="MIR"
27944..28238 /rpt_family="L1"
28703..28739 /rpt_family="Alu"
28734..28842 /rpt_family="AT_rich"
29236..29591 /rpt_family="Alu"
29590..30113 /rpt_family="MER1_type"
29726..29763 /rpt_family="AT_rich"
30116..30543 /rpt_family="L1"
30528..30570 /rpt_family="L1"
30544..30851 /rpt_family="Alu"
30852..31448 /rpt_family="L1"
31468..34303 /rpt_family="L1"
31752..31774 /rpt_family="L1"
34392..34840 /rpt_family="L2"
35098..35432 /rpt_family="L1"
35433..35743 /rpt_family="Alu"
35744..37636 /rpt_family="L1"
36078..36099 /rpt_family="L1"

Query Match 5.3%; Score 73; DB 9; Length 101034;
Best Local Similarity 44.3%; Pred.No. 0.028;
Matches 397; Conservative 0; Mismatches 490; Indels 10; Gaps 2;

QY 1 TAAATAATCTATACATTAAAAAATTCGATTTTAAATTTTAGAATTCATGATTTTATT 60
Db 77226 TTATATATAATATATATATATATATATATATATATATATATATATATATATATATA 77285

QY 61 TTTTTCACGAAATCCGTTAATATGTTTAAATATTTACCAACTPAATTTTAAATTTTA 120
Db 77286 TAAATATAATATACATAATATATATATATATATATATATATATATATATATATATA 77345

QY 121 TTTTAAAGCAATTAAGCATGTTTGTATAAAATATATATATTTTAAATACCTTTTCAAA 180
Db 77346 TTATATATAATATATATATATATATATATATATATATATATATATATATATATATT 77405

QY 181 GTATAAGTTGATGATGGCGTGGTAGATTATTTTAGTTCTAGGTTTCAATGCAAGTT 240
Db 77406 ATATATAATATAATGTTATATAATATATATATATATATATATATATATATATATAT 77465

QY 241 GGTTTAGACATTTAGCCCTATCTTTTCTTAAACCAAAATAAATGTAATGAAACCTT 300

```

```

Db 77466 ATATAATATAATTTATATATATATATATATATATATATATATATATATATATAT - 77524
QY 301 TAGGAAAAAAGAAATCAAAATGAAAACATCATCGGTGGAGTCGAGAGCCACACC 360
Db 77525 ----AATATAATATATATATATATATATATATATATATATATATATATATATATTA 77580
QY 361 CAGTGACCCCAACAATTTAAATTAAGAGTTTCTCTACAGTAAATCCGATACCTTTTTTA 420
Db 77581 TATTATATAATTTATATAATATAATATAATATAATATAATATAATATAATATAATTA 77640
QY 421 TTCAATACCTTTTCCACTTCTAAATCTTCGAGAGTTTGCACCGTTAACTAAATTAAGTGT 480
Db 77641 TATAATATAATATGATTTATATAATATAATATAATATAATATAATATAATATAATTTGA 77695
QY 481 ATATCCAAACGGTCTCTAAAAAACTTGTGTACCGTGTCTACATTTCAACTTTCCGCACCC 540
Db 77696 TTAIGTATATATATATATATATATATATATATATATATATATATATATATATATATG 77755
QY 541 TGAAGCCGTTATGTTTAGGTTAGTGTGTTGCAACAGTTGAAGCGCATCACTCAGGAGGCT 600
Db 77756 TATATTATATATATATATATATATATATATATATATATATATATATATATATATATA 77815
QY 601 ACTTGGTCTTTCGCTTTTCGCTTTTGTTCATTTTTCACGTGATTTTGTGGTGAACACG 660
Db 77816 AATTATATATATATATATATATATATATATATATATATATATATATATATATATATATA 77875
QY 661 CGTACTTGAAAACTTATTATAAATTACATAATTTTATAAGTTTCACCTTCTTATATAACT 720
Db 77876 TATAATATAAGTATATATATATATATATATATATATATATATAAAGTATATATATAT 77935
QY 721 CATATAATATATAGGTTTAGAATGCCAATTTTAAATAAAGAAATAAAAAATAAATAGA 780
Db 77936 TATATAATATAAAGTATATATATATATATATATATATATATATAAAGTATATATATGT 77995
QY 781 ATAAATCGAAAAAATGAAATGTAAAAAATTTTCAGGGGACAAATAAAATATCAAGTCT 840
Db 77996 ATATTATATAATATAAAGTATATATATATATATATATATATATATATATAAAGTATATT 78055
QY 841 ATTATTTAAATTTTCCATTAGAATTCATTTTCCCTTAGTTAAATPATGAGCTAGCCAGT 897
Db 78056 ATGTATATTATATATAATATAAAGTATATATATATATATATATATATATATATAAAGT 78112

```

Search completed: February 24, 2004, 09:43:19
Job time : 6109.31 secs

BEST AVAILABLE COPY